Spatial Demography (2026) 14:4
https://doi.org/10.1007/s40980-026-00160-y

RESEARCH

# Intra-Urban Inequalities in a Southern European City: The Case of Palermo

**Giuliana La Mantia**[1] · **Vincenzo Giuseppe Genova**[1]

Received: 1 October 2025 / Accepted: 5 February 2026


## Abstract

This study examines intra-urban socio-economic inequalities in Palermo, analysing the degree of variation in household consumption expenditure across three geographical units: census tracts, First-Level Unit, and neighbourhoods. For each geographical unit, we present the distinct patterns of intra-urban inequalities identified and discuss them in the light of the socio-urban evolution of the city of Palermo. The analysis relies on the integration of two different data sources, the 2011 Census microdata and the 2019 Household Budget Survey, through a statistical matching technique. In this way, a synthetic dataset was obtained that includes information on household expenditure and their area of residence. A multilevel modelling approach is therefore used to exploit the hierarchical structure of our data, where households are grouped in nested territorial units. The results show that, even if most of the variation in consumption expenditure is due to household characteristics, significant territorial differences persist. The greatest between-area variation emerges at the census tract and neighbourhood levels, revealing patterns of macro- and micro-segregation.

**Keywords** Intra-urban inequalities · Multilevel modelling · Micro-segregation · Southern Europe · Statistical matching · Palermo city

✉ Vincenzo Giuseppe Genova
vincenzogiuseppe.genova@unipa.it

Giuliana La Mantia
giuliana.lamantia01@gmail.com

[1] Department of Economics, Business, and Statistics, University of Palermo, Palermo, Italy

Published online: 25 February 2026

## 1 Introduction

Today, more than half of the world's population lives in cities, a share expected to rise to 68% by 2050 (Kilroy, 2012). Urbanisation concentrates people, services, and economic activities, fostering economic growth, innovation, and social interaction (Glaeser, 2011). At the same time, it can intensify existing social inequalities and generate new spatial forms of exclusion. Urban inequalities have a spatial manifestation in the uneven distribution of social groups, commonly referred to as spatial segregation (Massey & Denton, 1988; Reardon & O'Sullivan, 2004; Nijman & Wei, 2020). Within cities, residential patterns are shaped by both individual preferences and constraints (Schelling, 1971; Vaughan & Arbaci, 2011). Evidence from several cities shows that socio-economic segregation is largely driven by high-income groups, who exhibit greater residential mobility and tend to settle in highly valued urban locations, such as central districts, coastal areas, or out-of-centre enclaves (OCDE, 2018; Van Ham et al., 2018, 2021). Conversely, low-income groups are often constrained to the cheapest dwellings and neighbourhoods (Tammaru et al., 2021a). Thus, concentrations of poverty are associated with lower levels of local services and infrastructure and may negatively affect individuals' health, educational outcomes and employment opportunities (Massey, 1990; Musterd, 2005; Benassi & De Falco, 2025).

Levels and patterns of segregation result from a combination of multiple factors operating at both the national and local levels. At the national level, scholars have emphasised the role of welfare regimes and their effectiveness, housing systems and policies, economic restructuring, and changes in labour markets (Sassen, 2001; Allen et al., 2004; Musterd, 2005; Van Ham et al., 2015). These studies reveal that spatial disparities tend to be more pronounced in contexts characterised by weaker welfare states, limited state control over land, and liberalisation of the housing market (Musterd et al., 2017; Comandon et al., 2018; Van Ham et al., 2021). At the local level, other studies have highlighted the importance of historical processes, institutional contexts, and local morphology in mediating the effects of broader structural forces on segregation (Marcuse, 2002; Maloutas, 2016; Arbaci, 2019; Van Ham et al., 2024).

Historically, European cities have benefited from relatively strong welfare systems (Esping-Andersen, 1990), which have contributed to lower levels of segregation. However, the recent increase in income inequality in Europe has raised concerns about growing spatial segregation (Tammaru et al., 2021b; Van Ham et al., 2021). These concerns are considered particularly relevant in southern European countries (S-Eu), where income inequality is relatively higher, and welfare systems are less effective in mitigating social stratification than those in northern Europe (Allen et al., 2004; Maloutas, 2016). Moreover, land regulation and planning have often been weak, with a long-standing tolerance of illegal construction, producing an urban structure that differs from the more planned and zoned models of Northern European cities. Indeed, S-Eu cities are characterised by dense historic centres, fragmented expansion, unplanned peripheralisation, and widespread informal housing provision (Malheiros, 2002; Allen et al., 2004; Arbaci, 2008; Maloutas, 2016; Feria-Toribio et al., 2024). In such settings, socio-economic disparities may emerge at very fine spatial scales, sometimes along single streets or even within the same building blocks (Maloutas & Spyrellis, 2016). Drawing on this literature, our study focuses on the

Southern European city of Palermo, one of the largest cities in southern Italy and the regional capital of Sicily. The city combines long-standing socio-economic disadvantage with a fragmented urban structure shaped by weak planning enforcement, informal construction, and uneven suburban growth. Indeed, comparative studies rank Palermo among the most segregated Italian cities, both in socio-economic and ethnic terms (Barbagli & Pisati, 2012; Benassi et al., 2020). Inequalities are therefore embedded at multiple spatial scales, making Palermo a compelling case for exploring how household characteristics and place-based factors interact to produce socio-economic divides. Despite growing attention to segregation in Southern Europe, detailed analyses of intra-urban inequalities in Palermo remain limited, particularly those that utilise household expenditure at a granular spatial scale. Indeed, in Italy, survey data on consumption are representative only at the regional level (Istat, 2015). In contrast, Census data are the only data source that provides detailed territorial information, but it does not include any expenditure or income information (Istat, 2023).

In this work, we provide an analysis of intra-urban socio-economic inequalities in Palermo through the lens of household expenditure at the small area level, integrating two data sources: i) the microdata from the 2011 Population and Housing Census and ii) the microdata from the 2019 Household Budget Survey. Using a statistical matching procedure (D'Orazio et al., 2006), we impute total monthly household consumption expenditure to the census households data using a set of common socio-demographic and housing variables. This approach produces a synthetic dataset in which detailed geographical information (census tracts) and expenditure measures are jointly observed. We then link additional geographical units – first-level units (FLUs) and historic neighbourhoods – to the census, which allows for a multiscale analysis of inequality. To account for the nested structure of the data, with households grouped within census tracts, FLUs, and neighbourhoods, we apply multilevel regression models (Goldstein, 2011). This framework allows us to distinguish inequalities related to household characteristics from those arising from area effects across territorial scales. In this work, we focus on three main research questions:

1. To what extent do household-level factors explain expenditure differences in Palermo?
2. How much variation occurs at different spatial scales (census tracts, FLU, neighbourhoods)?
3. What historical and spatial processes underlie these inequalities, and how do they relate to the socio-urban evolution of the city?

The remainder of the paper is organised as follows. Section 2 introduces the Palermo context in a comparative and historical perspective. Section 3 describes the data, the statistical matching approach, and the modelling strategy. Section 4 presents and discusses the results. Section 5 concludes.

## 2 Palermo Context

Palermo is the capital city of the Sicily region and the fifth most populous city in Italy. Table 2 presents a comparison between Palermo and other major Italian cities, based on various demographic and socio-economic Census indicators. In 2011, the city reported a particularly low incidence of foreign residents, at 29.9 per 1,000 inhabitants, compared to a national average of 67.8. The city also ranks below the capitals of central and northern Italy, in particular Rome (85.8) and Turin (69.7), which record values above the Italian average. On the contrary, Palermo's value is more in line with that of the other southern capital, Naples, which has an incidence of 33 foreign residents per 1000 inhabitants.

In terms of residential property, the average size of dwellings in Palermo is $102m^2$, slightly above the national average (around $100m^2$) but considerably larger than the other capitals considered (around $80 - 91m^2$). Naples, Turin and Palermo have the most obsolete housing stock among the cities compared, with an average age of dwellings of approximately 37 years for Naples and Turin and 35.2 years for Palermo, all above the national average of 30.1 years. Conversely, Rome and Milan have a relatively younger housing stock, at 33.1 and 29.4 years, respectively. The two southern capitals are also characterised by relatively high rates of housing in poor condition, 3.9% for Naples and 3.7% for Palermo. These values are more than double the national average (1.7%) and also exceed the figures for central and northern capitals. Palermo shows particularly high values for the urban housing potential indicator, calculated as the percentage of unoccupied dwellings out of the total number of dwellings, which stands at 14.5%, considerably higher than other capitals, although lower than the national average (20.3%). This highlights the underutilisation of the city's existing housing stock, combined with a relatively high number of buildings in poor condition.

Palermo is home to the regional government offices and one of the largest universities in southern Italy; however, the graduate rate is 4 percentage points below the national average and lower than all other capitals except Naples. In fact, the city struggles to retain its young, educated population and is characterised by a persistent emigration of graduates and skilled workers (Istat, 2013). Socio-economic indicators are significantly worse than the national level, such as early school leaving (25.8%), youth unemployment (64.4%), and the NEET rate (38.8%). Economically vulnerable households in Palermo (7.3%) are more than double the national rate (2.7%). These conditions reflect the structural weaknesses of the local labour market, characterised by relatively low overall and female employment rates, as well as a heavy dependence on bureaucratic sector (non-commercial tertiary employment).

In comparative analyses at the national and European level, Palermo often ranks among the Italian cities with the highest levels of socio-economic and ethnic segregation (Barbagli & Pisati, 2012; Benassi et al., 2020). Interestingly, this is the case despite the migrant population being relatively small, confirming what has been found in other contexts, namely, a negative correlation between the percentage of migrants and the level of segregation (Benassi et al., 2020). A detailed analysis of the segregation patterns of the main ethnic groups in Palermo is provided by Busetta et al. (2015). Their results show that spatial settlement patterns reveal significant contrasts between migrant groups. However, unlike cities in northern Europe, Palermo does not show a clear trend towards suburbanisation: foreign residents tend to concentrate in central or semi-central areas, mirroring the patterns found in other southern port cities (Arbaci, 2008).

Other studies have analysed socio-economic disparities in Palermo, focusing on specific neighbourhoods such as Borgo Vecchio, ZEN, and Albergheria. These areas represent the most marginalised parts of the city, characterised by poor housing conditions, a high concentration of vulnerable households, high levels of school dropouts, informal work, and poor public infrastructure (Guarrasi, 1978; Capursi & Giambalvo, 2006; Picone, 2016). More recent works also identify forms of socio-spatial exclusivity in the city, with the development of *gated communities* in the urban periphery, which reinforce models of separation of wealthy families from the rest of the city (Tulumello, 2017) (table 1).

From a historical perspective, the current inequalities in Palermo originate from long-term urban development, which has subsequently been intensified by post-war urbanisation, producing new forms of spatial disparities (Inzerillo, 1984; Chubb, 1982; Takeuchi, 1990).

The first expansion of Palermo beyond the city walls occurred in the late 18th century, when development shifted northward. Along the new roads, the emerging bourgeoisie and the nobility began to build their own residences. This trend continued in

**Table 1** Comparison between Palermo and some regional capital cities, based on demographic and socio-economic Census (2011) indicators

| Indicator | Palermo | Naples | Rome | Turin | Milan | Italy |
|---|---|---|---|---|---|---|
| Resident population (absolute value) | 657 561 | 962 003 | 2 617 175 | 872 367 | 1 256 211 | 59 433 744 |
| Population density (res/km$^2$) | 4094.6 | 8082.5 | 2033 | 6709.9 | 6914.7 | 196.8 |
| Foreign residents (per 1,000 italians) | 29.9 | 32.7 | 85.8 | 69.7 | 62.7 | 67.8 |
| Home ownership | 60.6 | 53.2 | 70.2 | 65.8 | 59.5 | 72.5 |
| Dwelling avg size (m$^2$) | 102.1 | 86.7 | 91 | 85.6 | 80.1 | 99.6 |
| Urban housing potential | 14.5 | 4 | 9.7 | 8.3 | 7.9 | 20.3 |
| Buildings avg age (year) | 35.2 | 37.2 | 33.1 | 36.8 | 29.4 | 30.1 |
| Poor condition buildings | 3.4 | 3.9 | 1.7 | 0.6 | 2.3 | 1.7 |
| Adults with higher qualification | 51.2 | 49.4 | 72.5 | 61.6 | 60.8 | 55.1 |
| Young with degree | 20.6 | 20.5 | 35.8 | 30.7 | 31.5 | 23.2 |
| Early school leavers | 25.8 | 28.1 | 9 | 13.7 | 14.9 | 15.6 |
| NEET 15-29 | 38.8 | 42 | 19 | 19.8 | 12.8 | 22.5 |
| Youth unemployment | 64.4 | 67.5 | 35.8 | 33.4 | 20.6 | 34.7 |
| Employment rate | 34.6 | 31.8 | 47.9 | 46 | 48.5 | 45 |
| Female employment rate | 25.5 | 22.4 | 41.8 | 40.1 | 39.8 | 36.1 |
| Agricultural employment | 3.5 | 2.7 | 1.3 | 1.3 | 1.1 | 5.5 |
| Industrial employment | 11.9 | 15.7 | 11.2 | 24.5 | 22.4 | 27.1 |
| Non-commercial tertiary employment | 67.1 | 63.8 | 70.8 | 58 | 57.4 | 48.6 |
| Commercial employment | 17.5 | 17.8 | 16.7 | 16.2 | 19.1 | 18.8 |
| Economically vulnerable families | 7.3 | 9.5 | 2.1 | 1.7 | 0.4 | 2.7 |

Source: Author's own elaboration on Census Istat data (Istat, 2023)

The indicators refer to % unless otherwise specified the 19th century with the construction of Via Libertá, a northward extension of Via Maqueda. The area developed into an upper-class residential neighbourhood (*quartieri signorili*), highlighting a pattern in the city in which elites moved away from the central and older neighbourhoods towards prestigious, newly built northern neighbourhoods (Takeuchi, 1990) (see Fig. 1 Politeama and Libertá neighbourhoods).

The biggest changes, however, occurred after World War II. Urbanisation in this period was shaped by three main factors: (i) the building boom promoted by national policies, in particular through the financing of large social housing projects (Chubb, 1982; ii) the appointment of mafia figures in the municipal administration, which allowed them to influence urban planning and construction (Chubb, 1982; Takeuchi, 1990); and (iii) the abandonment of the historic centre, which was not rebuilt after the war, in favor of suburban expansion (Inzerillo, 1984; Cannarozzo, 2000).

At the time, Palermo became a clear example of “urbanisation without industrialisation", i.e. rapid demographic growth without a corresponding development of the productive economy (Chubb, 1982). The construction sector replaced industry as the main driver of local growth, and residential construction became the main instrument of social mobility, especially for civil servants who left the countryside to take up administrative positions in Palermo, which assumed the role of regional capital (Chubb, 1982; Cancila, 2014; Pedone, 2019).

Social housing played a crucial role in the expansion of the city (Chubb, 1982; Pedone, 2019). However, instead of serving as a tool to reduce housing inequality,

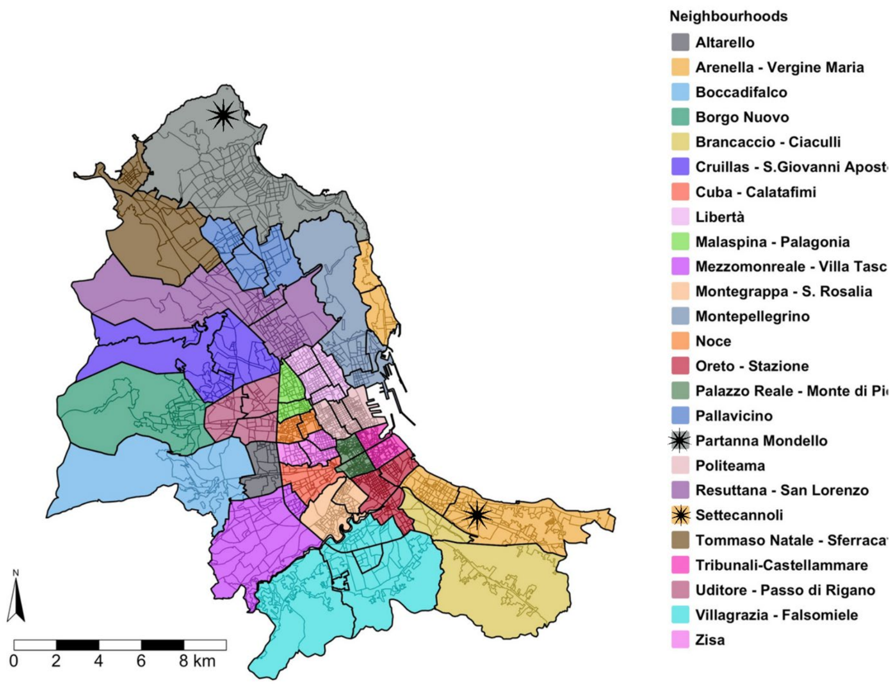



**Fig. 1** Repartition of the city of Palermo: census tract (gray borders), FLU (black borders), Neighbourhood (colors). *Source*: Author’s own elaboration on Istat data (Istat, 2001)

it mainly benefited property speculators and private builders at the expense of the public interest. In fact, public housing projects were located in the far north and far west of the city's planned urban area. Their distance from the established city made it more expensive to build the necessary infrastructure, the costs of which were covered by the public budget. This peripheral location was therefore bound to create opportunities for private speculation on the land in between, between the new neighbourhoods and the existing urban fabric, where multi-storey buildings for middle- and upper-class families rapidly sprang up. These intermediate areas correspond to the traditional northward expansion of upper-middle-class residential neighbourhoods. In addition, social housing was often distributed as political patronage to bureaucrats and trade unionists rather than being allocated to eligible assignees (Inzerillo, 1984; Badami, 2012).

Finally, the neglect of reconstructing the historic centre, which had been devastated by bombing during World War II, fostered depopulation of the areas and the formation of pockets of poverty and slums (Chubb, 1982). The 1968 earthquake, which rendered many buildings in the historic centre unsafe, further accelerated the exodus of residents from the inner city. The housing crisis and the misallocation of social housing dwellings produced waves of instability in the 1970 s and 1980 s, culminating in the occupations of new peripheral neighbourhoods such as the areas of San Filippo Neri (also known as Zona Espansione Nord – ZEN), San Giovanni Apostolo (also known as Centro Espansione Periferica – CEP), and Borgo Nuovo (see Fig. 1). These neighbourhoods remain marked by decay and marginalisation, consolidating new forms of social and spatial inequality (Chubb, 1982; Picone, 2016; Tulumello, 2017). This brief examination highlights how urban development in the city was characterised by the promotion of private interest and the tolerance of illegal construction, which has driven expansion in the absence of state control, resulting in a fragmented and small-scale urban growth (Chubb, 1982; Maloutas, 2016; Arbaci, 2019).

# 3 Data and Methods

## 3.1 Data

As mentioned in the introduction, this paper aims to examine intra-urban socio-economic inequalities in Palermo. To achieve this goal and overcome the limitations of a single data source, we integrate the Household Budget Survey (HBS, 2019) and Microdata from the 15th Census of Population and Housing (2011), through a Statistical Matching procedure, illustrated in Sect. 3.2.1.

### 3.1.1 Household Budget Survey (HBS)

The Household Budget Survey (HBS), conducted annually by ISTAT, allows for the analysis and monitoring of the evolution of household spending behavior according to their social, economic, and housing characteristics (Istat, 2015).

The statistical unit considered in the analysis is the household. The HBS collects detailed socio-economic variables from all the household members and also gathers data on housing characteristics. This set of socio-demographic and housing variable, called *core variables*, are widely included in other statistical surveys to facilitate statistical data integration and comparison across different sources (Istat, 2015, 2023). The survey primarily focuses on all expenditures incurred by households to purchase goods and services for consumption, while expenditures for purposes other than consumption are generally excluded (Istat, 2015). The variable we are interested in is total household consumption expenditure. It is important to point out that, although mortgage and rent payments are not included in the concept of consumption expenditure according to the ISTAT classification, this information is collected and, as it represents a substantial part of household housing costs, we include it in our total household expenditure variable. From the whole sample, we selected the observation units (households) from Sicily, since the survey provides representative estimates up to the regional level (Istat, 2015), which comprises 1 091 observations.

The HBS is an annual survey and the choice to use 2019 as a time reference is due to: (i) ensure the use of an ongoing survey; (ii) guarantee a sufficiently large Sicilian sub-sample (at least 1,000 units); (iii) avoid years affected by events that could affect consumption behaviour (such us pandemic or economic crisis).

### 3.1.2 The 15th Census of Population and Housing

The 15th Population and Housing Census, conducted in 2011, is one of the main sources of information on household characteristics and building attributes. It was the last ten-year Census carried out with exhaustive territorial coverage and a door-to-door questionnaire. The statistical unit considered in the analysis is the household and its associated dwelling. The Census collects socio-economic variables on households and their members and data on housing characteristics. Many of these variables are in line with those collected in the HBS, albeit with different coding structures. This alignment allows for matching procedures between the two data sources.

The 15th Census introduced several methodological innovations, paving the way for the integrated use of administrative sources and sample surveys in the permanent census. A key innovation was the adoption of a dual questionnaire, with *short* and *long* forms. The short form collected essential demographic, socio-economic, and housing information, while the long form included the full set of traditional census variables. To reduce the statistical burden on households and contain data collection costs, this strategy was implemented in municipalities with more than 20,000 inhabitants, where the long form was administered to a representative sample of households and the short form to the rest of the households (Istat, 2023). Thanks to an agreement with Istat, we were able to access and process the census microdata, which for privacy reasons are not available to the public. This data provides information on the census tracts in which each household resides. Census tracts are irregular polygons representing a small geographical area and constitute the most granular level at which census data has been provided. It is possible to construct higher-level geographical and administrative entities by aggregating census tracts.

In this study, we use three different geographical units: (i) census tracts (n=1,932); (ii) First-Level Unit (FLU, n=55); (iii) historic neighbourhoods (n=25). The FLU and historic neighbourhoods geographical units represent the old administrative areas of the city of Palermo. Figure 1 illustrates the geographical units present in our dataset. The census tracts are outlined by gray lines in the background, the FLU are identified by black borders, while the neighbourhoods are represented by distinct colours. Table 2 provides summary statistics on the distribution of households in the geographical units.

The final Census dataset comprises 89,229 observations (households), including information on household members and dwellings. Our selected sample includes (i) individuals living in households and residing in the municipality of Palermo and (ii) households that completed the *Long* questionnaire.

## 3.2 Methodology

### 3.2.1 Statistical Matching

Statistical matching (SM) encompasses a set of techniques for integrating different data sources to investigate relationships between variables that are not jointly observed in a single survey (D'Orazio et al., 2006; Rässler, 2012). Traditional SM methods rely on a variety of imputation models, such as regression imputation, hot-deck imputation, and predictive mean matching (D'Orazio et al., 2006; Lewaa et al., 2021).More recently, a growing body of literature has turned to *statistical learning* (SL) techniques which offer several advantages, overcoming some of the limitations of traditional methods (D'Orazio, 2019; Küntzler, 2025).

For instance, hot-deck imputation is a non-parametric method which involves grouping units into pools, based on matching variables, and imputing values from a *donor* to a *recipient*, either randomly within the same pool (random hot-deck) or based on a distance metric to find the most similar unit (nearest neighbour distance hot-deck) (D'Orazio et al., 2006). In practice, this requires researchers to manually select matching variables or define and compute distances between units. These steps are time-consuming and have a considerable impact on the final results (D'Orazio et al., 2019; Küntzler, 2025). Moreover, crossing a large number of matching variables can lead to empty or very small donor pools, compromising the reliability of the imputation (Spaziani et al., 2019). Similarly, defining a meaningful distance metric for the nearest neighbour hot-deck procedure is not straightforward when the set of matching variables includes both categorical and continuous variables (Spaziani et al., 2019; Küntzler, 2025). Furthermore, regression-based methods often rely on

**Table 2** Distribution of population across the geographical units. Census tract (n =1932), FLU (n=55), Neighbourhoods (n=25)

| Units | Min | 1st Qu | Median | Mean | 3rd Qu | Max | SD |
|---|---|---|---|---|---|---|---|
| Census tract | 10.0 | 18.8 | 34.0 | 46.2 | 59.0 | 631.0 | 42.0 |
| FLU | 247.0 | 1040.5 | 1516.0 | 1622.3 | 2139.5 | 3843.0 | 808.8 |
| Neighbourhood | 1072.0 | 2614.0 | 3418.0 | 3569.2 | 4744.0 | 6275.0 | 1532.5 |

Source: Author's own elaboration on Istat data (Istat, 2023) assumptions of linearity and may fail to capture non-linear relationships and higher-order interactions between covariates (Leulescu & Agafitei, 2013; Küntzler, 2025).

In contrast, statistical learning algorithms can directly accommodate mixed-type predictors, bypassing the need to select matching variables or define distance metrics manually. Furthermore, algorithms such as tree-based models are able to automatically capture non-linear relationships and complex interactions between covariates, which are often missed by linear regression models (Leulescu & Agafitei, 2013; Donatiello et al., 2014; Spaziani et al., 2019; D'Orazio, 2019; Küntzler, 2025). Despite these advantages, SL approaches can be computationally intensive and require hyperparameter tuning to avoid overfitting. Overall, the use of statistical learning algorithms in the literature on statistical matching has increased in recent years (Leulescu & Agafitei, 2013; D'Orazio, 2019; Spaziani et al., 2019; Li et al., 2021; Küntzler, 2025). Among others, D'Orazio et al. (2019) evaluates the performance of different SL algorithms by comparing them with hot-deck approaches. The results are promising, suggesting that statistical learning methods can improve predictive accuracy compared to traditional methods, although some approaches may be less reliable in preserving the marginal distribution of the target variable. Among the algorithms considered, boosting-based procedures offer one of the best compromises between the different assessment criteria. Similarly, Spaziani et al. (2019) proposes a mixed SM strategy in which statistical learning is used in the first phase, followed by a traditional hot-deck phase. In the comparison between alternative algorithms, boosting procedures again achieve particularly good results, both in terms of predictive accuracy and preservation of the association structure in the synthetic dataset. In fact, modern machine learning algorithms, particularly ensemble methods, have demonstrated considerable potential in prediction tasks (James et al., 2013).

Drawing on this literature, we implement a statistical matching procedure that employs a *Gradient Boosting Machine* (GBM) as the predictive algorithm. Therefore, unlike algorithms that rely on a single predictive model, GBM builds an ensemble of weak learners and combines them into a strong predictive model (James et al., 2013). GBM begins with a weak learner, in this case, a decision tree, which serves as the initial model. The residuals are then calculated, and a new decision tree is fitted to them. This process is repeated iteratively, with each new tree addressing the prediction errors from the previous stage (Boehmke & Greenwell, 2019). Thus, the core idea behind boosting is to sequentially improve the model by fitting new predictors to the residual errors of the previous ones.

Thus, by leveraging a set of common variables, including socio-demographic and housing characteristics (see Table 3), we can construct a synthetic dataset that enables us to explore the relationships between the household residence area, collected in the Censud data, and the total household consumption expenditure, collected in HBS data. This synthetic dataset enables the analysis of intra-urban inequality at small-area levels while preserving confidentiality, since the data are derived from imputation rather than direct observation (D'Orazio et al., 2006). In our implementation, the Census dataset is set as the *recipient* dataset and the HBS dataset as the *donor* dataset. The synthetic dataset is obtained by imputing total household consumption expenditure into the Census data, using information provided by the HBS. The procedure consists of two main steps: the GBM is first trained on the HBS data to predict house-

**Table 3** Households distribution by *common variables* in the census microdata (2011) and the household budget survey (2019)

| Core variables (X) | Level | Census count | Census (%) | HBS count | HBS (%) |
|---|---|---|---|---|---|
| HRP age group | 19 - 34 | 10767 | 11.66 | 44.11 | 9.12 |
| | 35 - 54 | 35704 | 38.66 | 171.82 | 35.52 |
| | 55 - 64 | 17374 | 18.81 | 96.38 | 19.92 |
| | 65 and over | 28517 | 30.88 | 171.48 | 35.44 |
| HRP nationality | Italian | 86980 | 94.17 | 455.15 | 94.08 |
| | Foreigner | 5382 | 5.83 | 28.65 | 5.92 |
| Household type | Single-person household | 24289 | 26.30 | 138.70 | 28.67 |
| | Couple without children | 16107 | 17.44 | 89.02 | 18.40 |
| | Couple with children | 37918 | 41.05 | 192.98 | 39.89 |
| | Single parent | 10078 | 10.91 | 34.14 | 7.06 |
| | Other | 3970 | 4.30 | 28.95 | 5.98 |
| Ownership status | Ownership | 52904 | 57.28 | 340.87 | 70.46 |
| | Rental | 28888 | 31.28 | 70.78 | 14.63 |
| | Other | 10570 | 11.44 | 72.14 | 14.91 |
| HRP education level | No education | 6701 | 7.26 | 45.98 | 9.50 |
| | Elementary or middle school | 46497 | 50.34 | 244.99 | 50.64 |
| | High school diploma | 24782 | 26.83 | 138.48 | 28.62 |
| | Tertiary education or higher | 14382 | 15.57 | 54.36 | 11.24 |
| HRP employment status | Employed | 42084 | 45.56 | 201.20 | 41.59 |
| | Unemployed | 4908 | 5.31 | 52.69 | 10.89 |
| | Seeking first job | 2855 | 3.09 | 7.21 | 1.49 |
| | Homemaker | 6768 | 7.33 | 69.64 | 14.39 |
| | Retired | 27147 | 29.39 | 132.23 | 27.33 |
| | Other condition | 8600 | 9.31 | 20.82 | 4.30 |
| Construction period | Before 1946 | 17269 | 18.70 | 53.38 | 11.03 |
| | 1946-1960 | 14671 | 15.88 | 52.55 | 10.86 |
| | 1961-1970 | 24188 | 26.19 | 88.62 | 18.32 |
| | 1971-1980 | 19150 | 20.73 | 112.55 | 23.26 |
| | 1981-1990 | 11913 | 12.90 | 103.63 | 21.42 |
| | 1991-2000 | 4049 | 4.38 | 36.33 | 7.51 |
| | After 2000 | 1122 | 1.21 | 36.74 | 7.59 |
| Surface area | Up to 70 | 23461 | 25.40 | 102.61 | 21.21 |
| | 71-90 | 23709 | 25.67 | 151.91 | 31.40 |
| | 91-110 | 22538 | 24.40 | 117.81 | 24.35 |
| | >110 | 22654 | 24.53 | 111.47 | 23.04 |
| Heating | Yes | 48917 | 52.96 | 276.39 | 57.13 |
| | No | 43445 | 47.04 | 207.41 | 42.87 |
| Air conditioning | Yes | 31033 | 33.60 | 239.47 | 49.50 |
| | No | 61329 | 66.40 | 244.33 | 50.50 |

Source: Author's own elaboration on Istat data (Istat, 2015, 2023)

hold consumption expenditure, and the fitted model is then applied to the Census data to impute the total household consumption expenditure.

A crucial assumption of statistical matching is the conditional independence assumption (CIA), which requires that the two variables of interest, observed in separate data sources, are conditionally independent given a set of common covariates. CIA allows us to solve the identification problem, whereby the conditional association between two variables that are never observed together cannot be estimated from the available data (Rässler, 2004; D'Orazio et al., 2006). Indeed, if the explanatory power of the common variables is high compared to the target variables, such that conditional independence is already established, the inference on the effectively unobserved association is valid (Rässler, 2004). In this way, the identification problem can be solved by decomposing the joint distribution into smaller estimation problems that can be solved with the observed data (D'Orazio et al., 2006). However, CIA is seldom valid and cannot be verified with the available data. The validity of the CIA becomes more plausible if the set of common variables includes strong proxies for expenditure and residential location, as this suggests that they capture most of the information explaining their relationship (Donatiello et al., 2016). In our case, we assume that the association between household consumption expenditure and the area of residence would be negligible once conditioned on the set of shared covariates.

Another assumption of statistical matching is that the two data sets refer to the same population or a subset thereof, and this can be verified by comparing the distribution of the set of common variables (Donatiello et al., 2014). Looking at Table 3, the distribution of common variables indicates overall comparability, although there are some discrepancies due to geographical and temporal differences between the two datasets. In particular, in the Census, household HRPs tend to be younger and more educated, and are more likely to live in rented accommodation than in the HBS, reflecting the fact that the Census refers to municipal rather than regional data. The HBS, on the other hand, shows a higher proportion of households living in newer dwellings equipped with heating/cooling systems, as well as a slightly higher proportion of individuals living alone, reflecting demographic changes over the last decade.

To evaluate the predictive performance of the GBM model, we implemented an out-of-sample evaluation based on a test set drawn from the donor survey (HBS). The model was trained on the remaining observations, and its accuracy was then evaluated on the test set using the coefficient of determination ($R^2$) and the root mean square error (*RMSE*). The resulting test set performance ($R^2 = 0.42$; *RMSE* =969) indicates that the set of common variables provides moderate predictive power for the target variable. Figure 2 shows the relative influence of the predictors, which summarises the relative contribution of each covariate to the predictive accuracy of the model (Ridgeway, 2007). Relative importance is a measure of the total improvement in the splitting criterion (the mean squared error) when a variable is used for splitting, calculated as an average across all trees (James et al., 2013). Our results show that the covariates with the greatest importance in predicting total expenditure are the head of household's level of education and his employment status. In addition, information about housing seems to contribute substantially, particularly the size of the dwelling, the presence of a heating system, and the period of construction of the building.

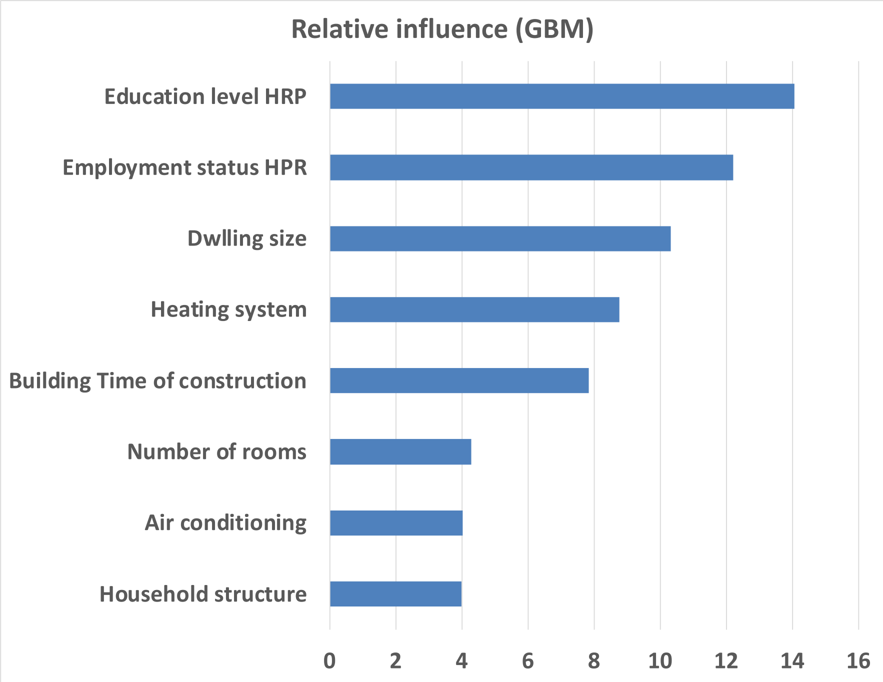



**Fig. 2** Relative influence of covariates. *Source*: Author elaboration

### 3.2.2 Multilevel Regression

To examine variation in household expenditure in Palermo, while controlling for household characteristics and accounting for the clustering of observations within geographical units, we employ a multilevel modelling approach. Multilevel models are well-suited for analysing nested data structures, where lower-level units are grouped within higher-level units, and their outcomes are influenced by these groupings (Goldstein, 2011). Ignoring such clustering may lead to biased standard errors and misleading inferences.

The hierarchical structure of our data (see Sect. 3.1.2) implies that households living in the same area are likely to be more similar to each other than to those in different neighbourhoods, introducing dependence between observations. The multilevel approach extends traditional regression techniques by introducing random effects that can vary at the group level. The inclusion of nested random intercepts enables us to capture group-specific deviations from the higher-level mean of the response variable. Multilevel models have been widely applied in the study of social inequalities and segregation, as they provide a robust statistical framework for analysing complex data structures (Goldstein & Noden, 2003; Leckie et al., 2012; Jones et al., 2015a, b; Fairbrother & Martin, 2013). A key advantage of this approach is the ability to quantify the extent of variation at each level of the hierarchy. Within this framework, the random effects variance can be interpreted as direct measures of inequality or segregation (Leckie et al., 2012). Larger variance estimates indicate greater dissimilarity between groups. After introducing explanatory variables in the model, the random effects variance captures unobserved heterogeneity that cannot be explained by observed covariates.

In our case, this framework enables us to quantify territorial inequalities that emerge at the micro (census tracts), meso (FLU), and macro (neighbourhoods) levels. For example, high residual variance at the census tract level would indicate strong contrasts between small adjacent areas. Conversely, higher variance at the neighbourhood level would suggest broader structural divisions within the urban fabric. We therefore specified, based on the likelihood ratio test with stepwise forward selection ($\chi^2 = 23.16$, $p < 0.001$) and the Akaike Information Criterion (AIC), a four-level radom four-level random intercepts model, in which the intercept varies across geographical units to reflect unobserved contextual influences. The explanatory variables are measured at the household level.

$$\log(y_{ijkm}) = \beta_0 + \sum_{c=1}^{C} \beta_{(c)} x_{(c)ijkm} + p_m + v_{km} + u_{jkm} + e_{ijkm}$$

$$\begin{aligned}
e_{ijkm} &\overset{iid}{\sim} \mathcal{N}(0, \sigma_e^2) && \text{Level-1 residual (Households)} \\
u_{jkm} &\overset{iid}{\sim} \mathcal{N}(0, \tau_u^2) && \text{Level-2 random effect (Census tracts)} \\
v_{km} &\overset{iid}{\sim} \mathcal{N}(0, \gamma_v^2) && \text{Level-3 random effect (FLU)} \\
p_m &\overset{iid}{\sim} \mathcal{N}(0, \delta_p^2) && \text{Level-4 random effect (Neighbourhoods)}
\end{aligned}$$

where $y_{ijkm}$ is the total equivalent household expenditure for household *i*, residing in census track *j*, within FLU *k*, and neighbourhood *m*. To correct for the typical right-skewed distribution of expenditure data, the response variable is log-transformed. Moreover, the Carbonaro equivalence scale (Carbonaro, 1985) is applied to total consumption expenditure to enable comparisons across household of different size. $\beta_0$ is the overall mean, $x_{(c)ijkm}$ is a set of household-level covariates, $p_m$, $v_{km}$, and $u_{jkm}$ are the random intercepts at the neighbourhood, FLU, and census tract levels, and $e_{ijkm}$ is the household-level residual. The random effects $p_m$, $v_{km}$, $u_{jkm}$ and the residuals $e_{ijkm}$ are assumed to be independent and normally distributed with mean zero and level-specific variances $(\delta_p^2, \gamma_v^2, \tau_u^2, \sigma_e^2)$ (Leckie, 2013). Specifically, $p_m$ represents the deviation of neighbourhood *m* from the overall mean, $v_{km}$ reflects the deviation of FLU *k* from the mean of its neighbourhood, and $u_{jkm}$ captures the deviation of census tract *j* from the mean of its FLU. The household residual $e_{ijkm}$ represents the deviation of household *i* from the mean of its census tract.

To interpret the variance components, we use the Variance Partitioning Coefficient (VPC), which measures the proportion of the total variance in the response variable that is attributable to differences between groups. It is derived from the decomposition of the total variance of the response variable as follows:

$$
\begin{aligned}
var(y_{ijkm}) &= var(\beta_0 + p_m + v_{km} + u_{jkm} + e_{ijkm}) \\
&= var(p_m) + var(v_{km}) + var(u_{jkm}) + var(e_{ijkm}) \\
&= \delta_p^2 + \gamma_v^2 + \tau_u^2 + \sigma_e^2
\end{aligned}
$$

By taking the ratio of each between-cluster variance component to the total variance of the response, we can assess the extent to which household expenditures vary between areas, as follows:

$$
VPC_p = \frac{\delta_p^2}{\delta_p^2+\gamma_v^2+\tau_u^2+\sigma_e^2}; \quad VPC_v = \frac{\gamma_v^2}{\delta_p^2+\gamma_v^2+\tau_u^2+\sigma_e^2}
$$
$$
VPC_u = \frac{\tau_u^2}{\delta_p^2+\gamma_v^2+\tau_u^2+\sigma_e^2}; \quad VPC_e = \frac{\sigma_e^2}{\delta_p^2+\gamma_v^2+\tau_u^2+\sigma_e^2}
$$

where $VPC_p$ represents the proportion of the total variance due to between-neighbourhood variation, $VPC_v$ the proportion due to between-FLU variation (within neighbourhoods), $VPC_u$ the proportion due to between-census tract variation (within FLU), and $VPC_e$ the proportion due to between-household variation (within census tracts). Then, for each geographical unit, posterior mean estimates are predicted. These estimates represent our main results and they are used to rank geographical units' level by their relative economic advantage or disadvantage, measured in terms of the deviation of the area's average expenditure from the higher-level average (Snijders & Bosker, 2011). It is worth noting that this ranking is based on shrunken residuals, which consider sampling variability and are more reliable for areas with smaller sample sizes, for which posterior estimates are shrunk towards 0 (Grilli & Rampichini, 2015; Raudenbush & Bryk, 2002). Finally, the residual spatial autocorrelation is tested using Moran's *I* and discussed in Sect. 4.

## 4 Results

The results summarised in Table 4 show the random effects for the four-level model without considering covariates (column 1) and after including covariates (column 2). The null model, or variance component model, represents how the variation in outcome expenditure is distributed across different levels (households and geographical units). It shows higher variation at both the micro-level (census tract) and macro-level (neighbourhood), and relatively small variation between meso-level (FLU). This pattern is partly consistent with that observed for the raw variance observed for each area, regardless of the hierarchical structure of the units. However, we can see that the response variation mostly reflects household differences within their census tract. This result is confirmed even after controlling for covariates. After introducing covariates at the household level, the residual variance at the household level decreases by 64%, while the variance between areas (between census tracts, FLU and neighbourhoods) is reduced by approximately 93%. This trend suggests that much of the variation observed between areas is attributable to household sorting and to the composition of the area in terms of household characteristics. However, the residual variance between areas is statistically significant, indicating that differences between areas persist even after accounting for household characteristics. This residual variation can be interpreted as evidence of *area-level effects* not captured by household

**Table 4** Random effects: Model 1: Variance Component Model; Model 2: Model with household level covariates. Relative variation of the variance components. Response variable: *log(total equivalent household expenditure)*

| | Model 1 | Model 2 | Relative variation | Observed variation |
|---|---|---|---|---|
| $\sigma^2$ | 0.0595 | 0.0212 | - 64 % | |
| $\tau_u^2$ | 0.0159 | 0.0012 | - 92 % | 0.04 |
| $\gamma_v^2$ | 0.0056 | 0.0004 | - 93 % | 0.02 |
| $\delta_p^2$ | 0.0128 | 0.0009 | - 93 % | 0.02 |
| VPC households | 63% | 89% | | |
| VPC census tracts | 17% | 5% | | |
| VPC FLU | 6% | 2% | | |
| VPC neighbourhood | 14% | 4% | | |
| Households | 89229 | 89229 | | |
| $R^2$ | 0.366 | 0.742 | | |
| Deviance | 6198.841 | −88303.856 | | |
| AIC | 1289839.438 | 1195571.051 | | |

Source: Author's own elaboration

characteristics, which may reflect unobserved contextual factors, such as local infrastructure, service accessibility, or building environment.

Table 5 presents the regression coefficient for the fixed effects of the model incorporating all selected covariates (Model 2 in Table 4). The intercept in this model represents the expected expenditure for the baseline household, i.e. a household characterised by a household reference person (HRP) with a high school diploma, living with a partner and children, residing in a home owned, built between 1961 and 1970 and equipped with both heating and air conditioning. Since the dependent variable is log-transformed expenditure, we report and interpret coefficients as approximate percentage changes. Among the regression coefficients, the education level of the HRP exhibits the highest coefficients and explains the greatest difference in household expenditure. Households with reference persons holding a degree or postgraduate qualification have a monthly expenditure that is 25.9% higher than those with a high school diploma. In contrast, those without any title or with only primary or middle school education have a monthly expenditure of 23% and 14% less than the baseline, respectively.

Single-person households exhibit the greatest variation in expenditure, with an estimated 27.1% higher monthly expenditure than couples with children, followed by couples without children (+13.9 %) and one-parent households (+10.5 %). The period of construction and the characteristics of the dwelling further influence the differences in expenditure. Households residing in pre-1946 buildings spend 17.3% less than those in dwellings built between 1961 and 1970, while those in newer constructions (post-1990) exhibit increased expenditure of 8.3% and 5.1% for dwellings built in 1991–2000 and after 2000, respectively. Households in homes without heating systems spend 12.2% less per month, and those without air conditioning spend 7.7% less, suggesting these appliances have a substantial impact on consumption patterns. Finally, households headed by foreign-born individuals spend 2% less than those with a native HRP, reflecting modest but significant differences.

**Table 5** Model 2 regression table. Response variable: *log(total equivalent household expenditure)*

| Predictors | | Estimates | % change | C.I. | p-value |
|---|---|---|---|---|---|
| | (Intercept) | 7.35 | | (7.33–7.36) | < 0.001 |
| Country of born of HRP: Italy | No native | −0.02 | −1.98 | (−0.02–0.01) | < 0.001 |
| Household structure: Couple with children | Couple with children | 0.13 | 13.88 | (0.13–0.13) | < 0.001 |
| | Single person household | 0.24 | 27.12 | (0.24–0.24) | < 0.001 |
| | Mono parent household | 0.10 | 10.52 | (0.10–0.10) | < 0.001 |
| | Others | 0.04 | 4.08 | (0.03–0.04) | < 0.001 |
| Education level of HRP: High school | No title | −0.26 | −22.89 | (−0.26–0.25) | < 0.001 |
| | Until middle school | −0.15 | −13.93 | (−0.15–0.14) | < 0.001 |
| | Degree or post degree | 0.23 | 25.86 | (0.23–0.23) | < 0.001 |
| Tenure: Homeownership | Rent | −0.08 | −7.69 | (−0.08–0.08) | < 0.001 |
| | Others | −0.05 | −4.88 | (−0.05–0.04) | < 0.001 |
| Heating system: yes | No | −0.13 | −12.19 | (−0.14–0.13) | < 0.001 |
| Air conditioning: yes | No | −0.08 | −7.69 | (−0.08–0.08) | < 0.001 |
| Time of construction: 1961–1970 | Before 1946 | −0.19 | −17.30 | (−0.20–0.19) | < 0.001 |
| | 1946–1960 | −0.10 | −9.52 | (−0.11–0.10) | < 0.001 |
| | 1971–1980 | −0.01 | −1.00 | (−0.01–0.00) | 0.004 |
| | 1981–1990 | 0.01 | 1.01 | (0.00–0.01) | 0.010 |
| | 1991–2000 | 0.08 | 8.33 | (0.07–0.08) | < 0.001 |
| | After 2000 | 0.05 | 5.13 | (0.04–0.06) | < 0.001 |

Source: Author's own elaboration

We now explore spatial inequality patterns through the analysis of posterior estimates. The results are presented in the next section, organised by our three geographical units: neighbourhood level, FLU level, and census tract.

### 4.1 Neighbourhood Level

The caterpillar plot in Fig. 3 displays the 25 neighborhood residuals along with their 95% confidence intervals. The zero line on the plot represents the overall mean expenditure across all neighbourhoods, i.e. at the city level. On the right side of the graph are the neighbourhoods with average expenditure above the city average, while on the left side are the neighbourhoods with average exependiture below the city average. Confidence intervals that do not overlap the zero line indicate significant deviations from the city average. Among the neighbourhoods with significant positive deviation, we identified a cluster of spatially contiguous neighbourhoods that expand in the norther part of the city. This cluster comprises the neighbourhoods: Libertá,which shows the highest residual at 0.06, followed by Politeama (0.05), Malaspina–Palagonia (0.04), and Resuttana–San Lorenzo (0.03). Currently, these areas are characterised by a concentration of affluent households. As mentioned in Sect. 2, this configuration stems from long-term historical dynamics that began with

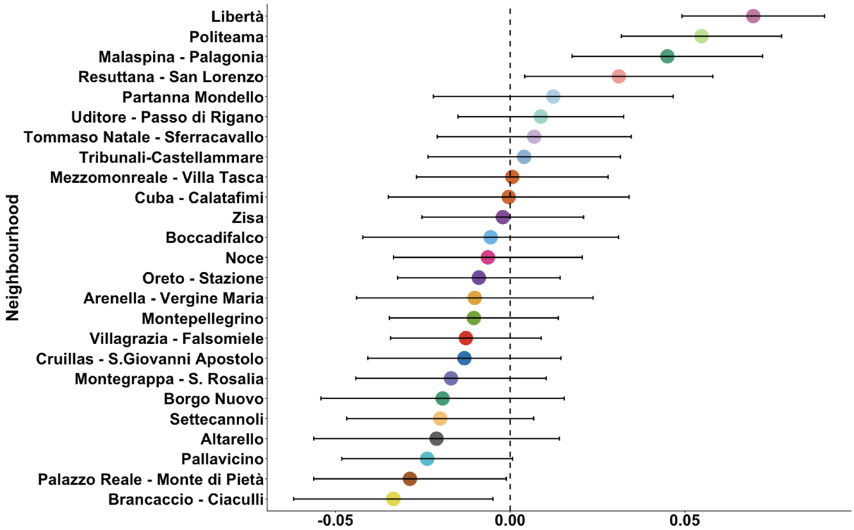



**Fig. 3** Caterpillar plot for neighbourhoods residuals and 95% confidence intervals for total equivalent consumption expenditure. On the right side of the graph are the neighbourhoods with average expenditure above the city average; on the left side are the neighbourhoods with average expenditure below the city average. The zero line on the plot represents the average expenditure at the city level. Confidence intervals that do not overlap the zero line indicate significant deviations from the city average. *Source*: Author's own elaboration

elite suburbanization from the city centre towards the northern new neighbourhoods since the late 18th century. This trend intensified after the Second World War, when historic villas and green areas were replaced by multi-storey buildings to accommodate the growing bureaucratic class (Inzerillo, 1984; Pedone, 2019). In this part of the city, residential building expansion was aggressive and took place without adequate infrastructure planning (Takeuchi, 1990). Moreover, many of the green areas originally included in the city's master plan were later built over, due to the ineffective enforcement of planning regulations (Inzerillo, 1984). Despite being the wealthiest parts of the city, these neighbourhoods still suffer from a structural shortage of public facilities and green spaces (Picone & Schilleci, 2012). Furthermore, this configuration reflects that of other port cities, where wealthy groups are concentrated in distinctive pericentral areas and along the coast (Arbaci, 2008).

By contrast, on the left side of the graph, two examples of disadvantaged areas can be identified: Palazzo Reale–Monte di Pietá (–0.03), one of the oldest neighbourhoods in the city, located in the inner area, and Brancaccio–Ciaculli (–0.035), one of the most recent neighbourhoods, located on the outskirts of the city.

The Palazzo Reale-Monte di Pietá stands out as one of the most disadvantaged neighbourhoods in Palermo. Despite its central location and the presence of important administrative buildings, the area is characterised by a high concentration of disadvantaged families. From a historical perspective, after the Second World War, the damage caused by bombing, combined with a long period of neglect, accelerated the deterioration of the historic centre, forcing many residents to move to the suburbs. The process of depopulation began after the war and continued until the early

1980 s (Chubb, 1982; Inzerillo, 1984). Over the last 40 years, the depopulation of the historic centre has slowed down, thanks to new migrants arriving in Palermo, who have often chosen to settle in the historic centre, where the poor condition of much of the housing stock makes it easier to find low-cost rental accommodation or occupy abandoned dwellings (Abbate, 2025). Since the mid-1990 s, there has also been a partial return of middle- and upper-class families to the city centre, and several redevelopment projects have been carried out (Abbate, 2025). However, these processes have only marginally affected the Albergheria neighbourhood, which has remained a place of physical and social degradation, as highlighted by a survey conducted in the early 2000 s by Capursi and Giambalvo (2006), which provides a detailed picture of social and housing conditions in Albergheria. The results reveal low levels of education among residents, early school leaving and employment patterns that show a strong dependence on informal work. In particular, housing conditions remain a central issue in Albergheria, where more than half of households (52.2%) live in rented accommodation and 14% live in allocated or occupied properties. Only 1% of dwellings have a heating system, and 24% of the buildings surveyed are classified as ruins, with foreign residents constituting the main occupants of these highly deteriorated buildings.

Moving towards the southern periphery, Brancaccio–Ciaculli emerges as one of the most disadvantaged neighbourhoods of Palermo (–0.035). As noted earlier, it was the last area to be incorporated into the urban fabric and historically consisted of agricultural settlements. Much of its built environment still reflects the features of a rural borough. Residential building development has concentrated along the roads closest to the city, particularly along Via Brancaccio, where large social housing blocks were constructed (Badami, 2012). The area also includes an industrial zone and a shopping centre; however, it remains both physically and socially separated from the rest of the city. The availability of services within Brancaccio-Ciaculli is inadequate, and the levels of green spaces, public amenities, and parking facilities fall below the city average (Picone & Schilleci, 2012). Although the opening of a tram line in 2015 improved transport connections, Brancaccio–Ciaculli continues to be classified among the city's most disadvantaged areas. Although our estimates are based on the 2011 Census micro-data, the main spatial patterns are consistent with more recent evidence from the ISTAT *Safety and State of Decay of Cities and their Suburbs* indicators, based on 2021 data (Istat, 2024). In particular, the northern neighbourhoods of Libertáá, Malaspina–Palagonia, and Resuttana–San Lorenzo appear to be the most attractive areas for highly educated individuals, with a share of residents with at least upper-secondary education above 80% (compared with 58% at the city level), a lower unemployment rate, and the highest home ownership rate. They also exhibit higher housing market values, on average around 1,600 €/m², compared to the city average of about 1,341 €/m². At the same time, these neighbourhoods are characterised by a more pronounced ageing profile: only about one in five residents is under the age of 24, and the old-age dependency ratio is among the highest in the city. On average, there are 240 elderly individuals for every 100 residents under the age of 14, compared to the city average of 159.

By contrast, neighbourhoods such as Palazzo Reale–Monte di Pietá and Brancaccio–Ciaculli show a younger demographic structure (around 30% under 24, compared with 24% citywide). However, they remain among the most disadvantaged areas. Unemployment is high (22.4% in both neighbourhoods, versus 17% at the city level), and the share of households with a reference person aged 65+ and no income source is around 10% almost double the city average of 5.8%. In addition, the NEET rate (15–29 years) is well above the city average (32.4%), reaching 45.3% in Brancaccio–Ciaculli and exceeding 50% in Palazzo Reale–Monte di Pietá. Housing conditions also differ markedly: in Palazzo Reale–Monte di Pietá, 61.5% of residents live in a dwelling they do not own (versus 34.5% citywide), and the share of improper dwellings is about five times the city average (1% vs 0.2%). Despite these disadvantages, housing values in Palazzo Reale–Monte di Pietá are only slightly below the city average (1,223 compared to 1,341), likely due to its central location, access to services, and appeal to foreign residents, with 200 foreign residents per 1,000 inhabitants. In contrast, Brancaccio–Ciaculli exhibit the lowest housing values, around 1,040. This can be attributed to its distance from the rest of the city, limited services and accessibility to services, and a lower presence of foreigners, with only 7 foreign residents per 1,000 inhabitants, which is lower than the city average at 38.4.

### 4.2 First Level Unit level

On a more detailed scale, now we examine the variation within neighbourhoods between FLU. Figure 4 shows the residuals for each FLU, measured against the average expenditure of the neighbourhood to which it belongs. In most cases, FLU values do not differ significantly from the average of the neighbourhood to which they belong. This confirms that FLU show greater internal heterogeneity and less dissimilarity between units, as also suggested by the VPC values (see Table 4).

However, some exceptions deserve further comment. In fact, within Politeama, belonging to the cluster of affluent neighbourhoods, there are strong expenditure disparities. On the one hand, the Croci-Ruggero Settimo FLU present a positive residual value of 0.04, reflecting above-average expenditure, while the Borgo Vecchio FLU has a residual value of −0.04, indicating a negative deviation.

Croci-Ruggero Settimo FLU belongs to the affluent areas of the city, developed during the city's first expansion northwards in the 18th century. At that time, Borgo Vecchio, founded in 1560 as a fishing village, was incorporated into the urban fabric. However, the area was not used by the middle class for the construction of new buildings as the proximity to the sea made it a unsafe place. Today, Borgo Vecchio still retains the topography of its original *borough* structure, with a dense network of narrow streets, small squares and irregular plots, which clearly differentiate it from the more modern urban fabric that surround it (Picone & Schilleci, 2013). In *La condizione marginale*, Guarrasi (1978) provides a detailed description of the area, highlighting the strong presence of working-class residents (57%) and the widespread availability of cheap housing in poor condition. Many of these dwellings lacked basic services, such as adequate sanitary facilities, and did not meet minimum living standards. According to the scholar, these precarious housing conditions, combined with the general state of degradation of the urban environment, contributed to the concentration of people without stable and remunerative employment in the area. The author also argues that the spatial closeness between Borgo Vecchio and the affluent

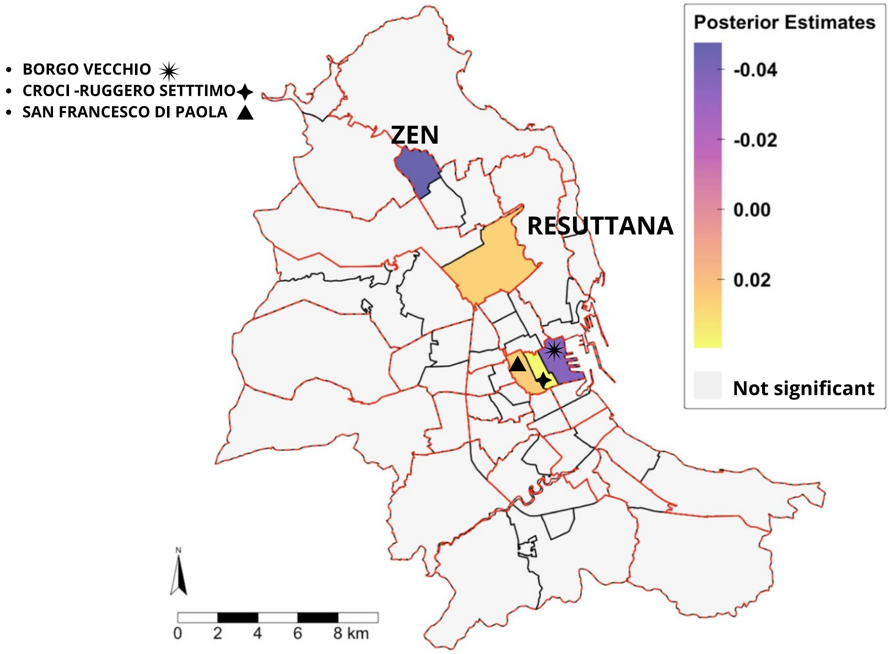



**Fig. 4** Map of First Level Unit residuals (within Neighbourhood, red bordes) for total equivalent household expenditure. Blue and purple census tracts represent areas with expenditure below the FLU average; yellow and orange census tracts represent areas with expenditure above the FLU average. *Source*: Author's own elaboration

part of the city has not reduced inequalities; rather, it has intensified social distance, accelerating the marginalisation of Borgo Vecchio and their inhabitants.

The FLU San Filippo Neri (or ZEN) presents a significant negative deviation, equal to −0.04, compared to the average of the Pallavicino neighbourhood, to which it belongs. This result confirms ZEN as one of the most disadvantaged areas of the city, as illustrated by a large literature on the subject (Bocchiaro & Tulumello, 2015; Fava, 2008; Picone, 2016). As mentioned in the Sect. 2, ZEN has been populated through several waves of illegal occupations in the 1970 s and in the 1980 s. The occupation of the houses created a dispute between the occupants and the municipality, which failed to complete the infrastructure works and connect the area to the rest of the city. The consequences of this negligence are still visible today. According to data from the 2011 Census, ZEN has one of the highest unemployment rates in the city: 51% overall, with youth unemployment reaching 79%, well above the city average of 64.4%. Moreover, 21% of households in the ZEN experience severe economic hardship, compared to the city average of 7.3%.

Today, ZEN is often described as an homogeneous ghetto, both because of its concentration of poverty and because of media narratives that reinforce its stigma (Bocchiaro & Tulumello, 2015). However, as Picone (2016) notes, ZEN is not homogeneous: its segregation is multi-scalar, with substantial internal contrasts between ZEN 1 and ZEN 2, further explored in the analysis of census tracts.

These patterns are also consistent with a recent small-area analysis by Carbonetti et al. (2025), who link Permanent Population and Housing Census data with administrative sources to construct a composite Socio-Economic Deprivation Index (SED-Index) for Palermo at the enumeration area (EA, i.e., census tract) level. They show that 12.2% of the 2 642 EAs have high deprivation (SED-Index > 110), and that about 1.2% fall into a very high deprivation class (SED-Index > 120). In their classification, ZEN clearly emerges as a highly deprived peripheral area, while Borgo Vecchio represents a deprived area in the city centre. ZEN displays a clear profile of cumulative disadvantage: a relatively young population (33.4% aged 0–24), very low employment among adults aged 35–64 (32.7%), and poor educational outcomes (76.2% of individuals aged 25–64 without upper secondary education; only 3.2% graduates), combined with a low presence of foreigners (3.1%). In contrast, Borgo Vecchio has a higher percentage of foreigners (22.1%) and a higher employment rate (49.5%), although no information is available on the types of work that are most common. In terms of educational outcomes, the prevalence of youth with a low qualification level and who are neither employed nor in training emerges in Borgo Vecchio. Housing disadvantage is visible in both FLUs. The share of households living in non-owner-occupied dwellings is very high, 76.1% in ZEN and about 60% in Borgo Vecchio, compared with 34.6% for Palermo overall, and the occupants-to-rooms ratio (0.5) is slightly above the city average (0.4).

### 4.3 Census Tract Level

The map in Fig. 5 illustrates the residuals at the census tract level, providing insight into the micro-level patterns of inequality in Palermo. The dashed red borders represent the FLU boundaries to stress that the average expenditure of each census tract is compared to the average expenditure of the FLU to which it belongs. Census tracts coloured in yellow and orange indicate higher-than-average expenditure within their FLU, whereas blue and purple census tracts represent lower-than-average expenditure. White census tracts, in contrast, denote areas where the average expenditure does not significantly differ from the FLUmean. The map highlights Palermo's fragmented urban landscape, where affluent and disadvantaged areas coexist in spatial proximity.

Zooming in on specific parts of the city in Fig. 6, three main forms of micro-level inequalities have been identified: (i) pockets of poverty; (ii) micro-polarisation along arterial roads; and (iii) Mediterranean-style gated communities.

Figure 6a shows a close-up of the historic centre. Within the FLU Palazzo Reale and Monte di Pietá, some census areas have significantly lower average expenditure than their FLU. These FLU belong to one of the most disadvantaged neighbourhoods in the city, and these blue and purple census areas can be categorised as pockets of extreme expenditure restriction.

It is interesting to note that these low-expenditure areas exist alongside relatively wealthy areas. In the same figure, in the Tribunali FLU, an example of micro-polarisation emerges along Via Lincoln. Here, the northern section, closer to Via Roma, shows higher expenditure (for example, Via Trento, Via Torino and Via Milano); while the southern part, which extends towards the seafront, includes two adjacent

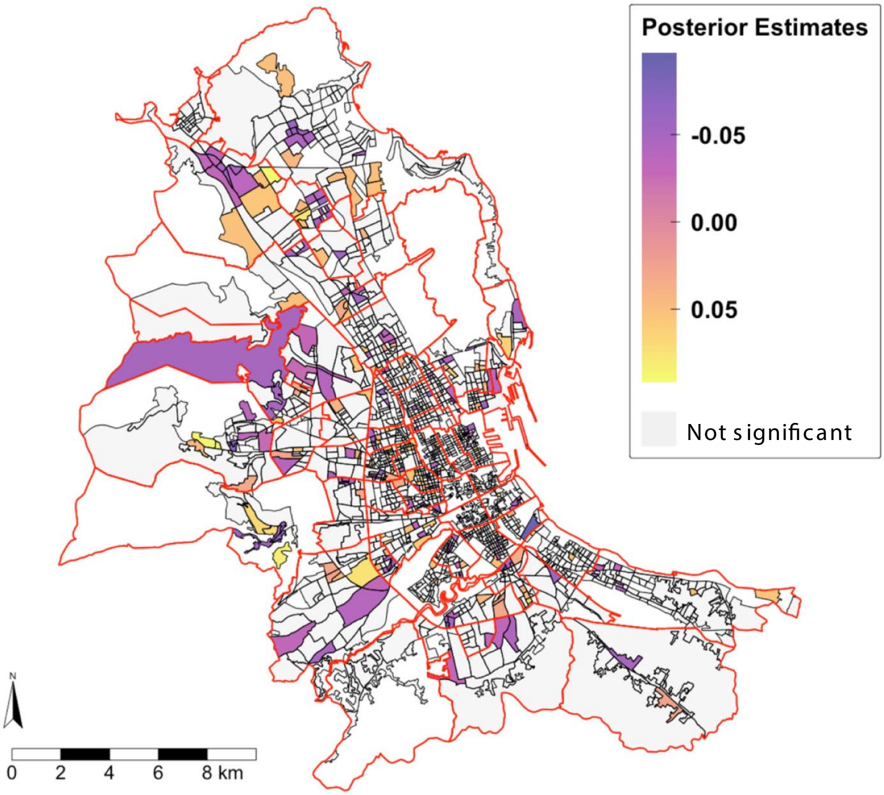
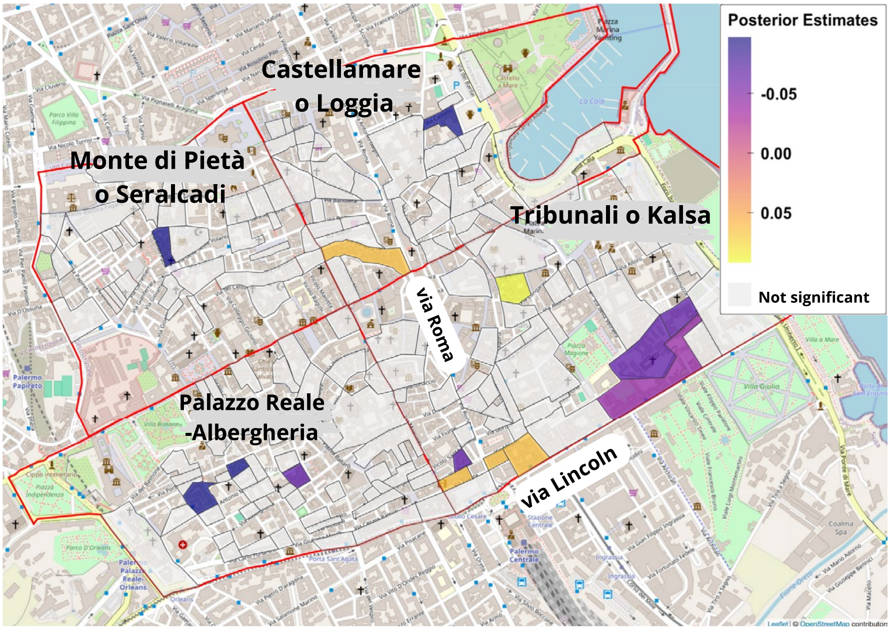


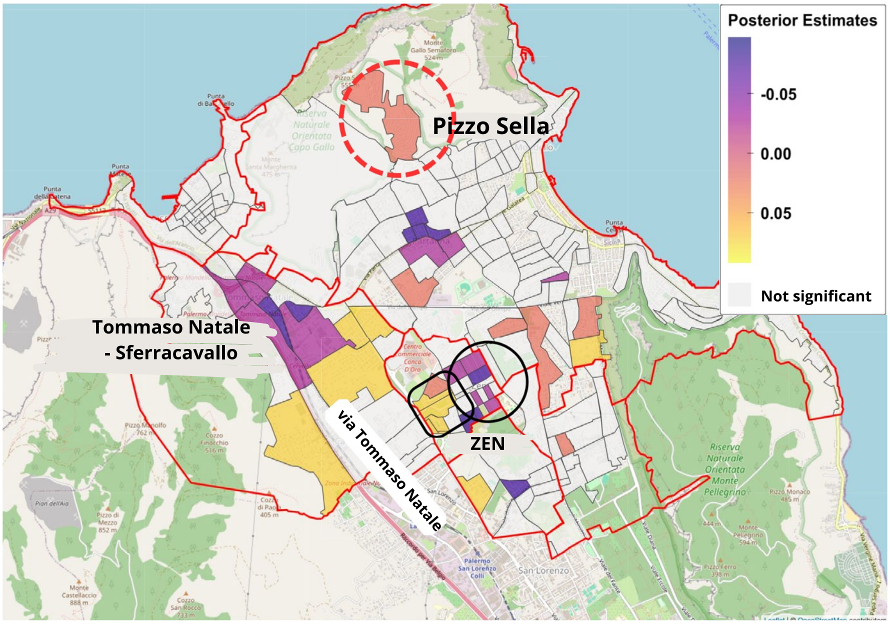

**Fig. 5** Map of census tract residuals (within FLU) for total equivalent household expenditure. Blue and purple census tracts represent areas with expenditure below the FLU average; yellow and orange census tracts represent areas with expenditure above the FLU average. *Source*: Author's own elaboration

census tracts characterised by significantly lower expenditure, especially around Piazza Magione and Piazza Kalsa. A similar spatial division appears in the FLU of Castellammare, where Via Venezia has above-average values, while the close area of Vucciria remains among the most economically disadvantaged areas of the city. These contrasts stem from urban transformations that took place during the Fascist era, when these areas were subjected to the cut-off by Via Roma. Along the new street, bourgeois buildings were constructed to attract middle-class families. In contrast, the neighbouring zones of Kalsa and Vucciria have retained their popular character, with people often living in crumbling structures and affected by informal tenancies.

Another case of micro-polarisation can be found along Via Tommaso Natale, in the far north of the city (Fig. 6b). On the south side of the road, high-expenditure areas feature a housing landscape characterised by detached houses and villas with private gardens and swimming pools, indicating the presence of upper-middle-class residents. To the north, beyond Via Nino Savarese, expenditure average levels decrease significantly. The area is home to the Marinella residential complex, a 1970 s project that envisaged the construction of a luxury residential complex. However, the project remained unfinished due to the bankruptcy of the company and the death of the




(a) Inner city




(b) North

**Fig. 6** Details of census tracts residuals (within FLU) for total household equivalent consumption expenditure. Blue and purple census tracts represent areas with expenditure below the FLU average; yellow and orange census tracts represent areas with expenditure above the FLU average. *Source*: Author’s own elaboration builder, who had links to the Mafia. Many unfinished units were subsequently occupied or assigned as social housing. Today, the area appears divided into two parts: on one side, there are well-maintained, gated apartment buildings where the residents are homeowners; on the other, illegal occupants and legitimate assignees are living in the most neglected apartment buildings (Picone & Schilleci, 2012; Tulumello, 2017).

Figure 6 balso illustrates the internal division within the ZEN. The census tracts of ZEN 1 (black square in the figure) record above-average expenditure within the FLU, while those of ZEN 2 (black circle in the figure) are significantly lower. This divergence reflects different histories of occupation and governance. In the 1970 s, ZEN 1 was partially occupied by organised movements or legitimate assignees demanding formal housing allocation (Lo Piccolo & Bonafede, 2007). In contrast, ZEN 2 developed through informal occupations linked to clientelist and criminal networks. Its *insulae* structure was never completed, and by the 1980 s, the area was already characterised by abandonment and Mafia control (Bocchiaro & Tulumello, 2015). Today, ZEN 1 benefits from modest integration thanks to transport links and infrastructure, while building block parts of ZEN 2 still lack basic infrastructure such as water, sewage, and electricity (Picone, 2016).

Finally, our findings highlight a form of micro-inequality that is not linked to social exclusion but to exclusivity: the emergence of gated communities. These are residential complexes with controlled access, often surrounded by automatic gates and equipped with shared spaces and facilities. They are designed to be spatially and socially separated from the surrounding urban fabric (Tulumello, 2015). As discussed by Tulumello (2017), gated communities in Palermo vary in size and services, but share some distinctive features. One common aspect is the privatisation of previously public roads, now reserved for the exclusive use of residents. Unlike North American models, surveillance is relatively limited and usually consists of guard posts and cameras at entry points rather than extensive high-security systems. The case identified in our results concerns the area of Pizzo Sella (red circle in panel 6b), part of the mountains surrounding the city to the north. The area has a higher average expenditure than the FLU to which it belongs (Partanna). Between 1978 and 1983, the area was parcelled out and more than a hundred licences were issued for the construction of dwellings. The area was therefore built up despite existing hydrogeological constraints and a subsequent demolition order. Even today, most of the buildings constructed are inhabited. The case of Pizzo Sella shows how weak land use regulations and ineffective enforcement of rules have fostered fragmented urban growth and socio-spatial clustering (Tulumello, 2015).

### 4.4 Spatial Autocorrelation

We test the residuals of the three geographical units for global and local spatial autocorrelation using Moran's $I$ and its local version (Anselin, 1995). For each geographical unit, we define a neighbour matrix using the queen contiguity criterion, whereby two areas are considered neighbours if they share either a border or a vertex (Lloyd, 2010).

Table 6 reports the values of the global Moran's $I$ and their significance levels under both the normality assumption and a random permutation approach. The null

**Table 6** Global Moran's *I* under normality assumptions and random permutation approach (nsim=999) for each geographical units

| Units | | Moran's I | Expectation | Variance | $Z(I)$ |
|---|---|---|---|---|---|
| Neighbourhood | Normality | −0.0679 | −0.0417 | 0.0143 | −0.219 |
| | Permutation | −0.0679 | −0.0453 | 0.0127 | −0.200 |
| First Level Unit | Normality | 0.0620 | −0.0185 | 0.00671 | 0.983 |
| | Permutation | 0.0620 | −0.0182 | 0.00672 | 0.978 |
| Census tract | Normality | 0.190 | −0.000518 | 0.000225 | 12.7 |
| | Permutation | 0.190 | −0.000695 | 0.000221 | 12.8 |

hypothesis is that there is no spatial autocorrelation. For neighbourhoods and FLU, we do not have evidence to reject the null hypothesis. By contrast, at the census-tract level, we detect a moderate positive global spatial autocorrelation, suggesting spatial clustering of residuals with similar values (either high or low).

Given this result, we further examine census tracts by decomposing the global statistic into its local components (local Moran's *I*) and testing their significance using the conditional randomisation approach proposed by Anselin (1995). This allows us to classify census tracts into the standard local Moran's *I* typology: high–high and low–low clusters (positive spatial autocorrelation) and high–low and low–high spatial outliers (negative spatial autocorrelation), where "high" and "low" are defined relative to the mean of the residuals (Anselin, 1995). Figure 7, which shows significant local clusters at the census tract level, provides further information on the micro-segregation patterns already highlighted by the residual analysis. Significant clusters primarily appear in the city's peripheral areas, particularly in the northern and western areas, where land use regulations and their enforcement have historically been weaker, and residential development has been strongly driven by private initiative. In these areas, we observe that census tracts in close spatial proximity belong to different cluster types, indicating that spatial proximity does not correspond to social proximity.

Finally, while local Moran's *I* provides further insight into the distribution of spatial inequality, inference based on the conditional randomisation approach involves a large number of simultaneous local tests, which raises a multiple-comparisons problem. After applying a False Discovery Rate (FDR) adjustment to the local p-values, none of the census tracts remain statistically significant, and the evidence of local spatial autocorrelation disappears, supporting the robustness of our estimates.

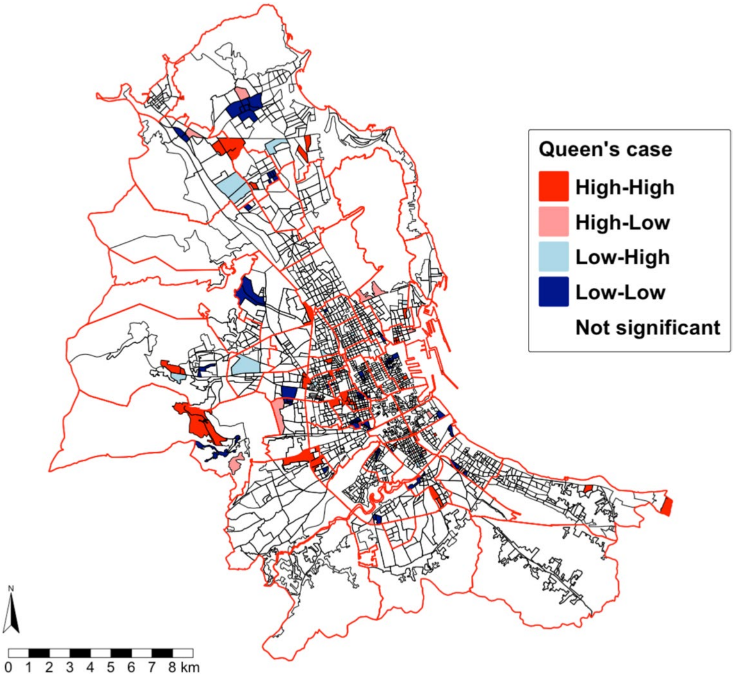



**Fig. 7** Local Moran's *I* significance cluster map at census tract level under conditional randomisation. *Source*: Author's own elaboration

## 5 Conclusion

This paper investigated intra-urban inequalities in the southern European city of Palermo by adopting a multilevel model approach. This framework enabled us to disentangle the extent of territorial variation in household expenditure across three geographical units: neighbourhoods, First-Level Units (FLUs), and census tracts. To conduct the analysis and overcome the lack of expenditure data at a detailed territorial level, we constructed a synthetic dataset that combined household expenditure information and household residence area. This was achieved through a statistical matching procedure that integrated the Household Budget Survey (2019), which contains expenditure information, with Census microdata (2011) that provides data on geographical units.

Our results reveal a multi-scalar structure of territorial inequalities in Palermo. The highest variation in household expenditure emerges between neighbourhoods and between census tracts, while variation between FLU is relatively small. We highlighted that many current patterns of inequality have long historical roots, dating back to the 18th century. At the same time, these disparities have been reinforced by post-war urbanisation dynamics, which have also produced new forms of spatial inequality.

At the neighbourhood level, the results show a clear pattern of macro-segregation: wealthy neighbourhoods clustered in the northern part of the city, while disadvantaged households are concentrated in the Palazzo Reale-Monte di Pietá area in the historic centre and in Brancaccio-Ciaculli on the southern outskirts. The highest degree of inequalities is found at the small-area level (census tracts). Especially, we identified three patterns of micro-segregation: (a) pockets of poverty; (b) micro-polarisation along main arteries; and (c) Mediterranean-style gated communities in the western and northern peripheral ring.

Palermo, therefore, combines a macro division between the north and the centre/south with a more scattered geography of inequality at the micro level. Pockets of poverty or enclaves of affluent families are not confined to the historic centre or a single suburb, but are scattered throughout the city. Furthermore, in some areas of the city, such as the far north, small areas of over-representation of affluent or disadvantaged families are spatially contiguous. Consequently, some parts of the city resemble a mosaic of social groups living close to each other, yet experiencing different housing conditions and lifestyles, which further exacerbates social inequalities.

This pattern aligns with observations in other southern European cities, where long-standing divisions have been redefined by post-war growth and, more recently, by pressures from the housing market (Arbaci, 2008; Malheiros et al., 2024; Maloutas & Karadimitriou, 2022).

In Athens, for instance, inequality is closely linked to post-war development, during which rapid population growth and a building boom led to the creation of a high-density residential city with limited services, especially in areas where the poorest groups reside (Karadimitriou et al., 2021). Since the late 1970 s, the suburbanisation of the elite from the centre has led to more socially homogeneous suburbs in the north-east and south-east, while the central areas have maintained a relatively mixed social composition (Maloutas & Spyrellis, 2000). A distinctive mechanism in Athens is the apartment-building system (*antiparochi*), which allows different social groups to live in the same building, resulting in a form of vertical segregation by floor level: wealthier households tend to occupy upper floors, while lower floors and basements are more accessible to less affluent residents, including many migrants arriving in the early 1990 s (Karadimitriou et al., 2021; Maloutas & Spyrellis, 2016). In this regard, Athens shows how social mix can coexist with inequality at a small scale, even within the same building. At the same time, the more general trends observed elsewhere, such as the strong centralisation of wealth and the peripheralisation of poverty, are not yet fully consolidated in Athens, where the contrast between a working-class west and a middle-class east remains observable (Maloutas & Botton, 2021; Maloutas & Spyrellis, 2000).

Conversely, the centre–periphery divide is more evident in Lisbon and Barcelona. In Lisbon, physical rehabilitation and regeneration of central areas have accelerated gentrification and contributed to the peripheralisation of working-class residents and labour migrants. In the Lisbon Metropolitan Area, top socio-professional groups (and EU migrants) are increasingly overrepresented in central and western Lisbon, as well as along the coastal strip towards Oeiras and Cascais, while lower-status groups concentrate more in the northern and eastern edges of the metropolis (Malheiros et al., 2024). Even though global indicators suggest a trend towards desegregation over recent decades, this may reflect a slow and fragmented process in which highly skilled and working-class blocks become spatially contiguous without real social interaction; in these cases, separation is detectable only at a micro scale, below the level of the *freguesia* (Malheiros et al., 2024). Barcelona exhibits a similar centre–periphery structure, with affluent groups concentrated in central and western districts, while vulnerability clusters in three main areas: the historic centre (Ciutat Vella, especially El Raval), the northeastern suburbs, and the southeastern neighbourhood of Besós-Maresme (Garcia-Almirall et al., 2021; Iglesias-Pascual et al., 2023). Here, the disadvantaged neighbourhoods reflect post-war migration and the construction of low-cost and often informal housing estates between the 1950 s and 1970 s, frequently in poorly connected and service-poor locations. At the same time, despite redevelopment programmes in areas such as El Raval, the residential composition of the area is characterised by extreme socio-economic fragility, which risks being further threatened by property speculation and the touristification of Barcelona's historic centre (Garcia-Almirall et al., 2021).

Overall, these comparisons reinforce the relevance of micro-segregation in Southern European cities. As defined by Maloutas and Karadimitriou (2022), micro-segregation refers to contexts where individuals living in spatial proximity occupy unequal positions due to socio-economic status or ethno-racial identity. From this perspective, social mixing at the neighbourhood level does not eliminate hierarchies, but can reproduce them within smaller units, even within the same building. Moreover, current urban processes, such as touristification and gentrification, pose new challenges for policies aimed at addressing urban poverty and spatial inequalities.

In terms of policy implication, our findings suggest that addressing socio-spatial inequalities in Palermo requires an integrated strategy combining area-based and individual-based interventions. In much of Western Europe, area-based policies often aims at social mixing, with housing provision and physical improvement used to attract middle-class residents to disadvantaged neighbourhoods (Maloutas & Karadimitriou, 2022; Van Ham et al., 2018). However, in the dense and compact cities of southern Europe, neighbourhoods with relatively low levels of segregation can conceal significant social inequalities on very small spatial scales (Arbaci, 2019). Furthermore, reducing spatial distance does not necessarily reduce social distance, and social integration policies can fail when they improve spatial integration without creating meaningful relationships or equitable access to services (Maloutas & Karadimitriou, 2022). These strategies also require substantial investment in housing, and if not carefully designed, can exacerbate the effects of gentrification and increase the risk of displacement for disadvantaged residents (Van Ham et al., 2018). This risk is particularly relevant in the historic centre of Palermo, where tourism-related investments and the rapid expansion of short-term rentals may increase pressure on still-affordable housing units (Crobe et al., 2025; Prestileo, 2021). This implies that housing and environmental regeneration should be accompanied by anti-displacement measures, such as regulating short-term rentals, providing social or affordable housing, and measures to protect vulnerable tenants (Karadimitriou et al., 2021; Maloutas & Karadimitriou, 2022). Furthermore, local policies should not only focus on central areas, but also on peripheral areas and informal settlements, where services, recreational facilities and basic infrastructure are still scarce. At the same time, data from international evaluations suggest that area-based investments can improve the physical environment but may have limited effects on individual outcomes if the structural factors of poverty are not addressed (Van Ham et al., 2018). Therefore, policies should also invest in opportunities for people by providing access to education and training programmes and facilitating job opportunities, especially for younger individuals in disadvantaged contexts. Finally, connectivity-based policies become crucial: improving transport accessibility can reduce barriers to reaching schools, services and workplaces across the city, particularly for residents of suburbs (Van Ham et al., 2018). Furthermore, this integrated approach also requires the involvement of resident groups and institutions at different levels, as well as the implementation of a local monitoring system through the creation of an urban observatory and the maintenance of a regularly updated database, which could support evidence-based targeting and help policymakers address socio-spatial inequalities (Karadimitriou et al., 2021). In this regard, our results can serve as a tool to inform policymakers in designing area-targeted policies that address social and territorial disparities.

### 5.1 Limits and Further Developments

Although our analysis has provided useful results, the study has some limitations that should be acknowledged. First, the statistical matching procedure relies on the conditional independence assumption (CIA), which is seldom valid and cannot be verified with the available data. If it does not hold, the matched data may lead to biased estimates and unreliable inferences (Donatiello et al., 2016). A second limitation concerns the temporal mismatch between the 2011 Census and the 2019 HBS. Statistical matching typically assumes that the sources refer to the same population, so that the distribution of common variables is similar. When reference years differ, comparability may weaken because socio-demographic and housing characteristics can change over time (see Table 3). In our case, the 2011 information may act as a lagged indicator of the residential context observed in 2019. This issue can be particularly relevant in urban settings, where mobility, redevelopment, and neighbourhood change can alter the composition of areas within a few years. For this reason, the spatial patterns identified here should be interpreted as reflecting the demographic and residential structure observed in 2011, rather than as an updated picture of 2019. Relatedly, the use of 2011 microdata also limits our ability to capture more recent urban dynamics. Over the last decade, changes such as infrastructure investments, demographic shifts, gentrification, and touristification may have reshaped Palermo's socio-spatial structure (Abbate, 2025; Crobe et al., 2025). This points to a broader constraint for both research and policy: the limited availability of up-to-date small-area microdata. Nevertheless, our main results are consistent with the 2024 ISTAT report on *Safety and State of Decay of Cities and their Suburbs* (Istat,, 2024) and with a more recent work of Carbonetti et al. (2025).

Finally, future research could advance in four main directions. First, further developments could adopt approaches that relax the CIA and provide probable bounds for the estimates of interest (D'Orazio et al., 2006; Küntzler, 2025; Rässler, 2004). Second, the multilevel model could be extended to include spatially structured random effects through a conditional autoregressive (CAR) specification, which introduces spatial dependence via the adjacency structure of areal units (Lee, 2013). Third, integrating richer and more recent sources, such as the Eurobarometer survey on quality of life in European cities (de Dominicis et al., 2023) and administrative registers, would help build a more timely and comprehensive picture of intra-urban inequalities. Finally, the approach presented here could be replicated in other cities to explore fine-grained patterns of urban inequality. This would enable comparative analyses between urban areas in southern Europe and cities belonging to different welfare and housing systems, allowing us to investigate how institutional contexts shape patterns of urban inequality and highlight commonalities and divergences in urban dynamics.

**Author Contributions** Giuliana La Mantia: Conceptualization, Methodology, Software, Formal analysis, Data Curation, Writing - Original Draft, Writing - Review & Editing. Vincenzo Giuseppe Genova: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding acquisition

**Funding** Open access funding provided by Università degli Studi di Palermo within the CRUI-CARE Agreement.

**Data Availability** The data used in this work comes from Istat and relates to the Census microdata (2011) and the Household Budget Survey (2019). The data processing was carried out at the Istat "Laboratorio per l'Analisi dei Dati ELEmentari" (ADELE) and in compliance with the rules on the protection of statistical confidentiality and personal data. The results and opinions expressed are the sole responsibility of the author and do not constitute official statistics. It should be noted that the analyses were conducted without using population weights.

## Declarations

**Conflict of interest** The authors declare no conflict of interest.



## References

Abbate, G. (2025). The Albergheria neighborhood in the historical center of Palermo between prospects for enhancing and conservation of identity. In R. Castanho, G. Hayder, & S. Ahmed (Eds.), *Urban identity explored: Architecture and arts in cities, advances in science, technology & innovation, vol CITAA 2023.* Springer.

Allen, J. Barlow, J. Leal, J., Barlow, J., & Maloutas, T. (2004). *Housing and welfare in Southern Europe*, vol 610. Wiley Online Library.

Anselin, L. (1995). Local indicators of spatial association—Lisa. *Geographical Analysis, 27*(2), 93–115.

Arbaci, S. (2008). (Re) viewing ethnic residential segregation in southern European cities: Housing and urban regimes as mechanisms of marginalisation. *Housing Studies, 23*(4), 589–613.

Arbaci, S. (2019). *Paradoxes of segregation: Housing systems, welfare regimes and ethnic residential change in Southern European cities*. Wiley.

Badami, A. (2012). I quartieri di edilizia residenziale pubblica a Palermo. In Picone M. & Schilleci F. (Eds) QU_ID. *Quartiere e Identità: Per una rilettura del decentramento a Palermo* (pp. 50–55). Alinea Editrice, Firenze.
Barbagli, M., & Pisati, M. (2012). *Dentro e fuori le mura. Città e gruppi sociali dal 1400 a oggi*. Il Mulino.
Benassi, F., & De Falco, A. (2025). Residential segregation and accessibility: Exploring inequalities in urban resources access among social groups. *Land, 14*(2), 429.
Benassi, F., Iglesias-Pascual, R., & Salvati, L. (2020). Residential segregation and social diversification: Exploring spatial settlement patterns of foreign population in southern European cities. *Habitat International, 101*, Article 102200.
Bocchiaro, G., & Tulumello, S. (2015). La violenza dello spazio allo zen di palermo. un'analisi critica sull'urbanistica come strumento di giustizia. *Archivio di S*tudi *Urbani e Regionali,* 73–94.
Boehmke, B., & Greenwell, B. M. (2019). *Hands-on machine learning with R*. Chapman and Hall/CRC.
Busetta, A., Mazza, A., & Stranges, M. (2015). Residential segregation of foreigners: An analysis of the Italian city of Palermo. *Genus, 71*(2–3), 177–198.
Cancila, O. (2014). *Palermo*. Laterza & Figli Spa.
Cannarozzo, T. (2000). Palermo: Mezzo secolo di trasformazioni. *Archivio di Studi Urbani e Regionali, 67*(2000), 1000–1039.
Capursi, V., & Giambalvo, O. (Eds.). (2006). *Al centro del margine: Standard di vita in un quartiere del centro storico di Palermo*. Franco Angeli.
Carbonaro, G. (1985). Nota sulle scale di equivalenza. In *Di indagine sulla povertà e sull'emarginazione C (ed) La povertà in Italia*. Istituto Poligrafico e Zecca dello Stato, Roma.
Carbonetti, G., Biasciucci, F., Cutillo, A., Mazziotta, M., Quondamstefano, V., Tamburrano, M. T., & Tronu, D. (2025). Measuring socio-economic deprivation at sub-municipal level through the integration of census and administrative data. *Rivista Italiana di Economia Demografia e Statistica*, 7–18.
Chubb, J. (1982). *Patronage, power and poverty in southern Italy: A tale of two cities*. Cambridge University Press.
Comandon, A., Daams, M., Garcia-Lopez, M., & Veneri, P. (2018). Divided cities: Understanding income segregation in OECD metropolitan areas. In *OECD (2018), dividedcities: Understanding intra-urban inequalities Paris* 23:25.
Crobe, S., Giubilaro, C., & Prestileo, F. (2025). Can culture save us? Rethinking culture-led touristification from Palermo (Italy). *Cities, 167*, Article 106331.
de Dominicis, L., Berlingieri, F., & d'Hombres, B. (2023). *Report on the quality of life in European cities*. Luxembourg: Publications Office of the European Union.
Donatiello, G., D'Orazio, M., Frattarola, D., et al. (2014). Statistical matching of income and consumption expenditures. *International Journal of Economic Sciences, 3*(3), 50.
Donatiello, G., D'Orazio, M., Frattarola, D., et al. (2016). The role of the conditional independence assumption in statistically matching income and consumption. *Statistical Journal of the IAOS, 32*(4), 667–675.
D'Orazio, M. (2019). Statistical learning in official statistics: The case of statistical matching. *Statistical Journal of the IAOS, 35*(3), 435–441.
D'Orazio, M., Di Zio, M., & Scanu, M. (2006). *Statistical matching: Theory and practice*. Wiley.
D'Orazio, M., Di Zio, M., & Scanu, M. (2019). Auxiliary variable selection in a statistical matching problem. In *Analysis of integrated data* (pp. 101–120). Chapman and Hall/CRC.
Esping-Andersen, G. (1990). *The three worlds of welfare capitalism*. Polity.
Fairbrother, M., & Martin, I. W. (2013). Does inequality erode social trust? Results from multilevel models of US States and counties. *Social Science Research, 42*(2), 347–360.
Fava, F. (2008). *Lo Zen di Palermo: Antropologia dell'esclusione* (1st edn.). FrancoAngeli.
Feria-Toribio, J. M., Iglesias-Pascual, R., & Benassi, F. (2024). Introduction. In *Socio-spatial dynamics in Mediterranean Europe: Exploring metropolitan structural processes and short-term change* (pp. 1–10). Springer.
Garcia-Almirall, P., Cornadó, C., & Vima-Grau, S. (2021). Residential vulnerability of Barcelona: Methodology integrating multi-criteria evaluation systems and geographic information systems. *Sustainability, 13*(24), 13659.
Glaeser, E. (2011). *Triumph of the city: How urban spaces make us human*. Pan Macmillan.
Goldstein, H. (2011). *Multilevel statistical models*. Wiley.
Goldstein, H., & Noden, P. (2003). Modelling social segregation. *Oxford Review of Education, 29*(2), 225–237.

Grilli, L., & Rampichini, C. (2015). Specification of random effects in multilevel models: A review. *Quality & Quantity, 49*, 967–976.
Guarrasi, V. (1978). *La condizione marginale, Prisma* (Vol. 11). Sellerio.
Iglesias-Pascual, R., Benassi, F., & Hurtado-Rodriguez, C. (2023). Social infrastructures and socio-economic vulnerability: A socio-territorial integration study in Spanish urban contexts. *Cities, 132*, Article 104109.
Inzerillo, S. M. (1984). *Urbanistica e società negli ultimi duecento anni a Palermo: crescita della città e politica amministrativa dalla ricostruzione al piano del 1962*. Istituto di Urbanistica e Pianificazione.
Istat. (2013). *Il benessere equo e sostenibile nelle città*. Istituto Nazionale di Statistica (ISTAT)
Istat. (2015). *La nuova indagine sulle spese per consumi in Italia*. Istituto Nazionale di Statistica (ISTAT)
Istat. (2016). *Descrizione dei dati geografici e delle variabili censuarie delle basi territoriali per i censimenti: anni 1991, 2001, 2011*. Istituto Nazionale di Statistica (ISTAT)
Istat. (2023). *Atti del 15° Censimento generale della popolazione e delle abitazioni. Volume 2: I metodi campionari e le tecniche di rilevazione*. Istituto Nazionale di Statistica (ISTAT)
Istat. (2024). *Commissione parlamentare di inchiestasulle condizioni di sicurezza e sullo stato di degrado delle città e delle loro periferie*. Istituto Nazionale di Statistica (ISTAT). Report Comune di Palermo.
James, G., Witten, D., Hastie, T., et al. (2013). *An introduction to statistical learning* (Vol. 112). Springer.
Jones, K., Johnston, R., Manley, D., et al. (2015). Ethnic residential segregation: A multilevel, multigroup, multiscale approach exemplified by London in 2011. *Demography, 52*(6), 1995–2019.
Jones, K., Owen, D., Johnston, R., et al. (2015). Modelling the occupational assimilation of immigrants by ancestry, age group and generational differences in Australia: A random effects approach to a large table of counts. *Quality & Quantity, 49*, 2595–2615.
Karadimitriou, N., Maloutas, T., & Arapoglou, V. P. (2021). Multiple deprivation and urban development in Athens, Greece: Spatial trends and the role of access to housing. *Land, 10*(3), 290.
Kilroy, A. (2012). *Intra-urban spatial inequality: Cities as " urban regions"*. World Bank.
Küntzler, T. (2025). *An automated machine learning pipeline for statistical matching* (p. 213). Foundations and Advances of Machine Learning in Official Statistics.
Leckie, G. (2013). *Three-level multilevel models - concepts*. LEMMA VLE Module 11.
Leckie, G., Pillinger, R., Jones, K., et al. (2012). Multilevel modeling of social segregation. *Journal of Educational and Behavioral Statistics, 37*(1), 3–30.
Lee, D. (2013). Carbayes: An r package for Bayesian spatial modeling with conditional autoregressive priors. *Journal of Statistical Software, 55*, 1–24.
Leulescu, A., & Agafitei, M. (2013). *Statistical matching: A model based approach for data integration*. Eurostat-Methodologies and Working papers, (pp. 10–12).
Lewaa, I., Hafez, M. S., & Ismail, M. A. (2021). Data integration using statistical matching techniques: A review. *Statistical Journal of the IAOS, 37*(4), 1391–1410.
Li, B., Chakraborty, S., Weindorf, D. C., et al. (2021). Data integration using model-based boosting. *SN Computer Science, 2*(5), 400.
Lloyd, C. (2010). *Spatial data analysis: An introduction for GIS users*. Oxford University Press.
Lo Piccolo, F., & Bonafede, G. (2007). Cronache zen: la questione abitativa tra assenza di politiche pubbliche, pratiche dal basso ed arte della negoziazione. *Archivio di Studi Urbani e Regionali, 90*, 1000–1020.
Malheiros, J. (2002). Ethni-cities: Residential patterns in the northern European and Mediterranean metropolises-implications for policy design. *International Journal of Population Geography, 8*(2), 107–134.
Malheiros, J., Silva, D. G., & Júnior, L. B. (2024). The contemporary geographies of urban inequality: Insights from Portugal and the Lisbon metropolitan area. In *Urban change in the Iberian Peninsula: A 2000–2030 perspective* (pp. 69–85). Springer.
Maloutas, T. (2016). Introduction: Residential segregation in context. In *Residential segregation in comparative perspective* (pp. 1–36). Routledge.
Maloutas, T., & Spyrellis, S. N. (2016). Vertical segregation: Mapping the vertical social stratification of residents in athenian apartment buildings. *Méditerranée Revue géographique des pays méditerranéens/Journal of Mediterranean Geography*, 27–36.
Maloutas, T., & Botton, H. (2021). Trends of social polarisation and segregation in Athens (1991–2011). *Social Inclusion, 9*(2), 117–128.
Maloutas, T., & Karadimitriou, N. (2022). Introduction to vertical cities: Urban micro-segregation, housing markets and social reproduction. In *Vertical cities* (pp. 1–21). Edward Elgar Publishing.

Maloutas, T., & Spyrellis, S. N. (2020). Segregation trends in Athens: The changing residential distribution of occupational categories during the 2000s. *Regional Studies*.

Marcuse, P. (2002). The partitioned city in history. In *Of states and cities: The partitioning of urban space of states and cities* (pp. 11–34). Oxford Academic.

Massey, D. S. (1990). American apartheid: Segregation and the making of the underclass. *American Journal of Sociology, 96*(2), 329–357.

Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces, 67*(2), 281–315.

Musterd, S. (2005). Social and ethnic segregation in Europe: Levels, causes, and effects. *Journal of Urban Affairs, 27*(3), 331–348.

Musterd, S., Marcińczak, S., Van Ham, M., et al. (2017). Socioeconomic segregation in European capital cities increasing separation between poor and rich. *Urban Geography, 38*(7), 1062–1083.

Nijman, J., & Wei, Y. D. (2020). Urban inequalities in the 21st century economy. *Applied Geography, 117*, Article 102188.

OCDE. (2018). *Divided cities: Understanding intra-urban inequalities*.

Pedone, F. (2019). *La città che non c'era: Lo sviluppo urbano di Palermo nel secondo dopoguerra, vol 6*. Istituto Poligrafico Europeo Casa Editrice.

Picone, M. (2016). Una segregazione paradossale e multi-scalare: il caso del quartiere zen di palermo. *Méditerranée Revue géographique des pays méditerranéens/Journal of Mediterranean Geography,* 37–46.

Picone, M., & Schilleci, F. (2012). *QU_ID QUartiere e IDentità: Per una rilettura del decentramento a Palermo*. Città e Territorio.

Picone, M., & Schilleci, F. (2013). A mosaic of suburbs: The historic boroughs of Palermo. *Journal of Planning History, 12*(4), 354–366.

Prestileo, F. (2021). Airbnb a Palermo: geografie, concentrazioni e attori. In Perrone, C., Masiani, B. e Tosi, F. (Eds.), *A cura di, Una geografia delle politiche urbane tra possesso e governo Sfide e opportunità nella transizione.* Dipartimento di Architettura dell'Università di Bologna.

Rässler, S. (2004). Data fusion: Identification problems, validity, and multiple imputation. *Austrian Journal of Statistics, 33*(1–2), 153–171.

Rässler, S. (2012). *Statistical matching: A frequentist theory, practical applications, and alternative Bayesian approaches,* vol 168. Springer.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods,* vol 1. Sage.

Reardon, S. F., & O'Sullivan, D. (2004). Measures of spatial segregation. *Sociological Methodology, 34*(1), 121–162.

Ridgeway, G. (2007). Generalized boosted models: A guide to the GBM package. *Update, 1*(1), 2007.

Sassen, S. (2001). *The global city: New York, London, Tokyo, updated* (2nd edn.). Princeton University Press.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology, 1*(2), 143–186.

Snijders, T. A., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.

Spaziani, M., Frattarola, D., & D'Orazio, M. (2019). *Integration of survey data in R based on machine learning*. Romanian Statistical Review.

Takeuchi, K. (1990). Urban politics and the urbanization process: A case study in Palermo. *Ekistics*, 16–21.

Tammaru, T., Knapp, D., Silm, S., et al. (2021). Spatial underpinnings of social inequalities: A vicious circles of segregation approach. *Social Inclusion, 9*(2), 65–76.

Tammaru, T., Sinitsyna, A., Akhavizadegan, A., et al. (2021). Income inequality and residential segregation in European cities. In *Urban inequality and segregation in Europe and China: Towards a new dialogue* (pp. 39–54). Springer.

Tulumello, S. (2015). From "spaces of fear" to "fearscapes" mapping for reframing theories about the spatialization of fear in urban space. *Space and Culture, 18*(3), 257–272.

Tulumello, S. (2017). Fearscapes: Urban space and the landscapes of fear. In *Fear, space and urban planning: A critical perspective from southern Europe* (pp. 55–85).

van Ham, M., Manley, D., & Tammaru, T. (2024). Geographies of socio-economic inequality. *Oxford Open Economics, 3*, i634–i641.

Van Ham, M., Marcińczak, S., Tammaru, T., et al. (2015). A multi-factor approach to understanding socio-economic segregation in European capital cities. In *Socio-economic segregation in European capital cities: East meets west.* Taylor & Francis.

Van Ham, M., Tammaru, T., & Janssen, H. (2018). A multi-level model of vicious circles of socio-economic segregation. In *Divided cities: understanding intra-urban inequalities* (pp. 135–153).

Van Ham, M., Tammaru, T., Ubarevičienė, R., & Janssen, H. (2021). Rising inequalities and a changing social geography of cities. An introduction to the global segregation book. In *Urban socio-economic segregation and income inequality: A global perspective* (pp. 3–26). The urban book series. Springer.

Vaughan, L., & Arbaci, S. (2011). The challenges of understanding urban segregation. *Built Environment, 37*(2), 128–138.