DEMOGRAPHIC RESEARCH

Research Article

Spatial attraction in migrants' settlement patterns in the city of Catania

Angelo Mazza

Antonio Punzo

Contents

Spatial attraction in migrants' settlement patterns in the city of Catania

Angelo Mazza[1]

Antonio Punzo[2]

Abstract

BACKGROUND
In broad terms, and apart from ethnic discriminatory rules enforced in some places and at some times, residential segregation may be ascribed both to economic inhomogeneities in the urban space (e.g., in the cost of rents, or in occupation opportunities) and to spatial attraction among individuals sharing the same group identity and culture.

OBJECTIVE
Traditional indices of spatial segregation do not distinguish between these two sources of clustering. Furthermore, they typically rely on census tracts, a scale that does not allow for fine-grained analysis. Also, the use of alternative zoning often leads to conflicting results. The aim of this paper is to measure spatial attraction among groups of foreign migrants in Catania (Italy) using individual household data.

METHODS
We apply a version of Ripley's K-function specially conceived for assessing spatial attraction while adjusting for the effects of spatial inhomogeneity. To avoid the risk of confounding the two sources of clustering, spatial inhomogeneity is estimated following a case-control approach.

RESULTS
Different parts of the city exhibit different suitabilities for migrants of different nationalities, with groups mainly involved in housekeeping and caregiving being more spread than the ones specialized in peddling and retailing. A significant spatial attraction has been found for Sri Lankan, Mauritians, Senegalese, and Chinese. Conversely, the settlement patterns of Tunisians and Moroccans comply with random allocation. These results

---

[1] Department of Economics and Business, University of Catania, Corso Italia, 55, 95128 Catania, Italy. Tel. (39) 0957537736, E-Mail: a.mazza@unict.it.
[2] Department of Economics and Business, University of Catania, Corso Italia, 55, 95128 Catania, Italy. Tel. (39) 0957537640, E-Mail: antonio.punzo@unict.it.

seem consistent with the hypothesis of a relevant correlation between chain migration and spatial attraction.

1. Introduction

Residential location influences individuals' proximity to important resources (such as schools, hospitals, child care facilities, labor markets, and employment opportunities) and to potential risks, including environmental threats and social hazards (such as exposure to crime and violence) (Reardon 2006). Furthermore, it impacts access to social networks and other forms of social capital; overall, it shapes human interaction and the demographic processes that originate from it, such as mortality, fertility and mobility (Almquist and Butts 2012).

A minority ethnic group is spatially clustered when the spatial arrangement of minority households departs from expectations based upon a random spatial allocation (Freeman, Pilger, and Alexander 1971).

In broad terms, and apart from ethnic discriminatory rules enforced by law or traditions in some places and at some times, we may distinguish between two sources of spatial clustering. One source is spatial inhomogeneity or apparent contagion. Typically, the different parts of a city exhibit large variations in the price of residential property, in the accessibility of low cost public infrastructures, and in the availability of certain types of jobs; these inhomogeneities may lead to a mostly economically induced segregation. As Schelling (1971) observes, ethnicity is often correlated with income, and income with residence; so even if residential choices were unconstrained by ethnic discrimination, the different ethnic groups would not be randomly distributed among residences.

The second source is spatial attraction or true contagion. Survey data on the ideal neighborhood composition for different ethnic groups in the USA, reported in Clark and Fossett (2008), show that all groups prefer living in areas where their group is a majority or near-majority. These preferences have complex origins and may reflect attachment to group identity and culture (e.g., language, religion, customs, etc.). Newly arrived minority migrants may benefit from positive spillovers in settling close to their compatriots, in terms of reciprocal acceptance, common language, and support. Transnational social networks play an important role in channeling arriving migrants into specific neighborhoods and also into particular occupations (Gelderblom and Adams 2006).

However, regardless of what the basis of the individual preferences for coethnic contact is, they produce identical patterns of residential segregation (Clark and Fossett 2008). The Schelling (1971) model provides an analysis of the implications of individual preferences and shows that when a household enters a neighborhood, that neighborhood becomes more attractive to members of the household's own group and less attractive to

members of other groups. In other words, the presence of a household in a given area increases the probability of others of the same group to locating nearby.

It is relevant in social research to be able to distinguish between these two sources of clustering. Whereas economic induced segregation might explain some initial degree of segregation and raises questions of social equity, the Schelling model highlights the importance of individually motivated segregation and posits that even mild preferences for living with similar neighbors carry the potential of being strong determinants for residential segregation (Clark and Fossett 2008).

The spatial distribution of households may be represented by a point pattern, i.e., a set of points in a map. Ripley's K-function (Ripley 1981) is widely used to detect clustering in point processes. The inhomogeneous K-function is a version of Ripley's K-function conceived for assessing the effects of spatial attraction (or inhibition), while adjusting for the effects of spatial inhomogeneity. In other words, this approach allows us to distinguish between the two sources of clustering, by assessing clustering above and beyond that due to apparent contagion.

Traditional studies of segregation rely on data aggregated by areal units, typically census tracts. However, analysis at the scale of the tract is unable to take into account geographic proximity at the level of the individual household and social scientists have begun to recognize the importance of the individual level analysis (Kumar 2007; Almquist and Butts 2012; Matthews and Parker 2013; Kumar, White, and Singh 2013). Furthermore, alternative zoning or a different scale often yields different results. This "modifiable area unit problem" is a well-known issue (Openshaw 1984, Wong 2009). In contrast, the K-function provides an explicitly spatial approach and an egocentric measure of clustering that takes into account questions of proximity as well as scale.

Although information on individual locations are often not available, for reasons of both privacy and administrative costs, simulation techniques may be used. Almquist and Butts (2012) propose point process models that can be used to effectively simulate the location of individual households. Also Kumar, White, and Singh (2013) use census tract data to simulate point locations to compute segregation indexes at finer spatial resolution.

In this paper we propose the use of the inhomogeneous K-function (Baddeley, Møller, and Waagepetersen 2000) to study spatial attraction in the household locations of some groups of foreign migrants in the city of Catania (Italy). The paper is organized as follows. Section 2 illustrates the methodological framework and in Section 3 a short outline on foreign migration inflows in the city of Catania is given. In Section 4 spatial inhomogeneity is estimated following a case-control approach; observed explanatory variables related to population density and rent cost per square meter are used. Then, the inhomogeneous K-function, is computed. Conclusions are given in Section 5.

2. Methodology

2.1 Spatial inhomogeneity and spatial dependence

The spatial distribution of households may be represented by a point pattern, i.e. a set of points in a map, where each point corresponds to the location of a single event. The use of the term "event" has become standard in spatial analysis, as a means of distinguishing the location of an observation (a household location in our case) from any other arbitrary location within the study region R.

The homogeneous Poisson process (HPP) is the simplest theoretical model of spatial point pattern. In this model, the number of events $N(U)$ in any region $U \subseteq R$ follows a Poisson distribution with mean

$$E[N(U)] = |U|\lambda$$

proportional to the area $|U|$ of U, where λ denotes the intensity of the process, which is assumed to be constant over R. Furthermore, in HPPs the number of events in non-overlapping sub-regions are independent random variables, thus there are no interactions among the events.

In applications, the HPP is often used as a benchmark model; the logic is that if a point pattern is completely random, then there is nothing "interesting" happening; in other words, the HPP is used as the "null model" of a statistical analysis (Baddeley 2010). However, population and its associated socio-economic and demographic attributes vary significantly across the geographic space and the constant intensity of the HPP is not suitable for representing the spatial distribution of the human phenomena (Kumar, White, and Singh 2013).

The inhomogeneous point process (IPP) is a generalization of the HPP obtained by replacing the constant intensity λ with a spatially varying intensity function $\lambda(u)$; so, we have

$$E[N(U)] = \int_U \lambda(u)du.$$

An IPP typically exhibits more variation in the location of events than the HPP does; clustered patterns occur, with regions where intensities are higher, receiving a higher number of events.

2.2 The inhomogeneous K-function

Ripley's K-function, usually denoted with $K(d)$, is widely used to detect clustering (or inhibition) in point processes with constant intensity; at every spatial distance d, $\lambda K(d)$ is the expected number of additional points of the process **X** located in a circle of radius

d surrounding an arbitrary event x. The K-function is commonly used in testing for HPP, and its asymptotic distributions under simple point process models are known (Heinrich 1991; Ripley 1981).

Baddeley, Møller, and Waagepetersen (2000) generalized Ripley's K-function to non-homogeneous point processes, by weighting each event x by $w = 1/\lambda(x)$. The inhomogeneous K-function is defined as

$$K_{\text{inhom}}(d) = E\left[\sum_{x \in \mathbf{X}} \frac{1}{\lambda(x)} \mathbf{1}(0 < \|u - x\| \le d) \,\middle|\, u \in \mathbf{X}\right], \tag{1}$$

where $\|u - x\|$ is the Euclidean distance between events u and x and $\mathbf{1}(\cdot)$ is the indicator function. Note that if $\lambda(u)$ is constant across space, $K_{\text{inhom}}(d)$ reduces to the usual $K(d)$ function.

For estimating K_{inhom}, Baddeley and Turner (2000) proposed the estimator

$$\widehat{K}_{\text{inhom}}(d) = \frac{1}{|R|} \sum_i \sum_{j \ne i} \frac{\mathbf{1}(\|x_i - x_j\| \le d)}{\widehat{\lambda}(x_i)\widehat{\lambda}(x_j)} e(x_i, x_j; d), \tag{2}$$

where $e(x_i, x_j; d)$ is an edge correction weight (Ripley 1991, p. 22), and $\widehat{\lambda}(u)$ is an estimate of the intensity function $\lambda(u)$.

2.3 Confounding between spatial inhomogeneity and spatial attraction

In real cases, clusters are most often generated by the joint action of both spatial inhomogeneity and spatial attraction. To distinguish between these two sources of clustering is challenging, as many spatial processes are "equifinal" and hence unidentifiable (Harvey 1966), i.e., one realization of a point process may be consistent with underlying processes involving clustering due to either spatial inhomogeneity or spatial attraction, as demonstrated in Bartlett (1963).

The risk of confounding the two sources of clustering is a fundamental limitation on the scope of statistical inference from a spatial point pattern, assuming we have access to only a single realization of the underlying process (Diggle et al. 2007; Baddeley 2010; Baddeley, Rubak, and Turner 2015) .

Diggle et al. 2007 outline three possible solutions. One is to assume a parametric model for the intensity, as in Møller and Waagepetersen (2003). Another is to assume that the effects of spatial inhomogeneity operate on a larger spatial scale than the effects of spatial attraction, thus allowing separation of the two. A third solution is applicable in case-control studies. In these studies, data consist of a realization of two spatial point processes, one representing *cases* of a condition of interest and the other representing

controls drawn at random from the population at risk. Assuming that the size of the population at risk tends to infinity and the sampling fraction to zero, the controls constitute a realization of an IPP with intensity $\lambda^*(u)$; the cases form a second, independent, point process – which may or may not be a Poisson process – with intensity $\lambda(u)$. The question of interest is whether the cases form an IPP with intensity proportional to that of the controls, i.e., $\lambda(u) = \rho\lambda^*(u)$, or whether they exhibit additional spatial structure. Diggle et al. (2007) propose that the ratio between the intensity functions may be modeled to depend on a vector of m spatially referenced covariates $\boldsymbol{z}(u) = (z_1(u), \ldots, z_m(u))'$, i.e.

$$\lambda(u) = \lambda^*(u)\, f\left(\boldsymbol{z}(u); \boldsymbol{\theta}\right), \tag{3}$$

where $f(\cdot)$ is any nonnegative function, either nonlinear parametric with parameters $\boldsymbol{\theta}$ as in (3), or nonparametrically specified. Note that within this approach, the estimate of $\lambda(u)$ depends on the case locations only through the values $\boldsymbol{z}(u)$ of the covariates at those locations, a feature that is critical in separating heterogeneity from clustering of the case point process (Diggle et al. 2007).

2.4 Inferential framework

Generally, only $\widehat{K}_{\text{inhom}}(d)$, the estimated counterpart of $K_{\text{inhom}}(d)$, is known and, to test the usual null hypothesis of absence of spatial interaction, a proper inferential framework needs to be introduced.

Because the exact distribution of $\widehat{K}_{\text{inhom}}$ is unknown, to obtain the distributional information needed, in the following we will consider Monte Carlo simulations from the null hypothesis. Roughly speaking, given n replications of the fitted IPP, a different value, say $\widehat{K}^j_{\text{inhom}}(d)$, $j = 1, \ldots, n$, may be computed for each replication. Pointwise confidence envelopes

$$L(d) = \widehat{K}^{(k)}_{\text{inhom}}(d)$$
$$U(d) = \widehat{K}^{(n-k)}_{\text{inhom}}(d)$$

of size $\alpha = 2k/(n+1)$ are obtained, where $\widehat{K}^{(k)}_{\text{inhom}}(d)$ denotes the k-th order statistic. Observed values of $\widehat{K}_{\text{inhom}}(d)$ falling, for some value of d, outside the envelopes indicate a significant departure from the null hypothesis at level α.

3. Foreign migration in the city of Catania

Towards the end of the 1970s, Catania started to become the destination of migration inflows coming from the poorest countries, with immigrants usually looking for domes-

tic work, catering, peddling and low qualification services. Migrants' work fulfilled a structural demand for cheap, flexible, and often off-the-books unprotected labor, a type of work that natives tend to avoid, even if unemployed.

These flows originate from all continents. Table 1 reports the distribution by nationality of households of foreign immigrants in the city of Catania (source: Catania Population Register Office on August 30, 2005). According to the Italian laws regulating cities' population registers (art. 4 of Presidential Decree 223/1989), a household is a group of people linked by marriage, affinity, adoption, guardianship or affective bonds who live together. Single person households are allowed.
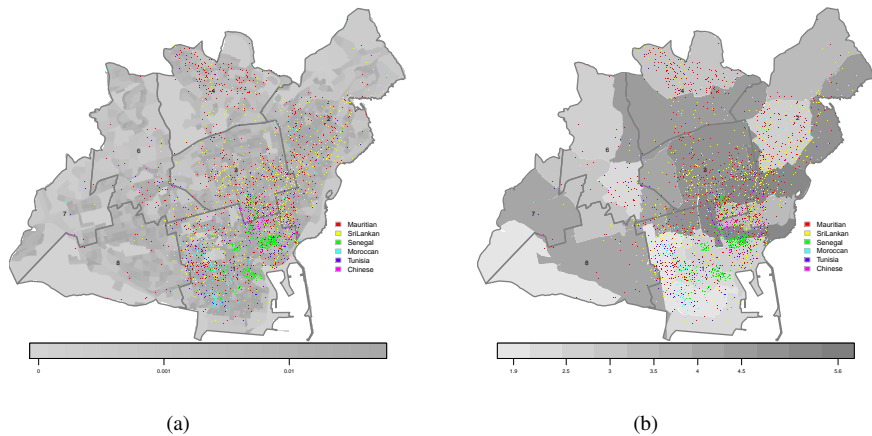
Table 1: **Households of foreign immigrants in the city of Catania on August 30, 2005 (Catania Population Register)**

Nationality	Foreign households in Catania households	%
Mauritian	1,073	24.06
Sri Lankan	811	18.18
Senegalese	469	10.52
Moroccan	188	4.22
Tunisian	183	4.10
Chinese	157	3.52
Greek	153	3.43
Bangladeshi	146	3.27
Colombian	95	2.13
Philippine	92	2.06
Romanian	78	1.75
Indian	65	1.46
German	61	1.37
French	55	1.23
Polish	50	1.12
British	49	1.10
Yugoslav	45	1.01
Pakistani	39	0.87
Ukraine	38	0.85
Maltese	37	0.83
Other	576	12.91
All countries	4,460	100.00

Previous studies highlighted the development of an ethnic work specialization (Licciardello and Damigella 2011, Altavilla and Mazza 2012, Busetta, Mazza, and Stranges 2016). Mauritians, Sri Lankans, Colombians, Indians, and Filipinos are mainly involved in housework within Italian's households. Romanians, as well as Ukrainian, Philippine, and Polish, are characterized by a high incidence of women, who are mainly care-givers for the elderly. Other groups, coming from Maghreb and Sub-Saharan Africa, espe-

cially males from Tunisia, Morocco, and Senegal perceive housework as "female" and degrading, and prefer to work as street vendors. They sell different types of – most often fake branded – goods, such as handbags, sunglasses, jewelry, and other inexpensive seasonal objects, along streets in the central shopping district, at open-air markets, at major road intersections and, in summer, along the nearby beaches. Finally, Chinese are mostly employed in import-export business and retailing. Such specializations, given the different logistic needs related to the occupational typologies and given the different integration possibilities, have favored the differentiation of the spatial distribution (Altavilla and Mazza 2012).

Figure 1:   **Immigrants' households locations**

(a)   (b)

Note: Foreign immigrants' households locations for the top six nationalities in the city of Catania on August 30, 2005 (source: Catania Population Register). Ordinal numbers identify municipalities. In the background: (a) Smooth pycnophylactic interpolation (Tobler 1979) of population counts for census tracts (2001 General Population Census, http://dawinci.istat.it). (b) Minimum rent cost per square meter for private residential properties (Italian Revenue Agency - OMI database, http://www.agenziaentrate.gov.it/geopoi/omi/).

In the following, we will focus on the households spatial patterns originated by immigrants afferent to the top six nationalities of Table 1. These patterns are shown in the maps in Figure 1. Note that the area of the city is subdivided in ten municipalities, identified by ordinal numbers. The 5th, the 9th and the 10th municipalities, not involved in the immigration process, are not shown. As we can see, the southern and western municipalities are the least influenced by the immigration process. Particularly affected by the migratory phenomenon is, instead, the 1st municipality, which, characterized by

a heterogeneous urban fabric, comprises the city's historic and political-administrative center and hosts the central business district as well as the largest outdoor market of the city, where native Italians and immigrants share the same space to sell their goods. The remaining areas, in the north portion of the 1st Municipality, for the most part host the middle or middle-high classes; they are mainly buildings erected in the 1950s and 1960s, characterized by decent multi-level apartment blocks and by higher building densities.

4. Migrant settlement patterns in the city of Catania

In this section we estimate the intensity functions $\lambda_i(u)$, with $i = 1, \ldots, p$, related to immigrants of the $p = 6$ most numerous nationalities. Then, the inhomogeneous K-function is used to describe and compare their settlement patterns.

Data at hand come from the office of the Catania population register, as recorded on August 30, 2005. The analysis includes households of immigrants and native-born children/grandchildren of immigrants who at birth only acquired the (foreign) citizenship of their parents. Register data have been integrated within a geographical information system (GIS), and all the residential addresses have been geocoded. For the estimation of $\lambda_i(u)$, other geographically referenced covariates have been used, such as census tracts data from the 2001 General Population Census and data on rent cost per square meter for private residential properties, collected for the different areas of the city by the Italian Revenue Agency (OMI database).

4.1 Estimate of spatial intensities

Different areas of the city have different suitabilities for different socio-economic groups of migrants, with the availability of a job being the single most important factor. In Section 3 we mentioned that a work specialization among migrants of different nationalities has developed, so that the different logistic needs led to a differentiation of the settlement patterns. Here we consider two occupationally related groups of nationalities: "domestic workers" and "vendors". The first group includes houseworkers and care givers and is made up of immigrants originating from Mauritius, Sri Lanka, Colombia, Philippines, Romany, India, Poland, and Ukraine. The second group includes street vendors and retailers and is made up of migrants from Senegal, Morocco, Tunisia, China, and Bangladesh. Figure 1 shows that whereas common workplaces for domestic workers are found in the wealthier and often peripheral residential neighborhoods, vendors benefit from locations that are close to the area surrounding the central business district. We deem it reasonable to assume that the environmental constraints within each occupational group are similar.

Here we will follow the case-control approach described in Section 2.3. Cases will

be, in turn, households belonging to each nationality i that we are researching, while the controls will be all households belonging to the nationalities of the same occupational group of i but with the exception of i.

Our data are not strictly case–control data; however, in epidemiological studies, the limited availability of data for non diseased persons often leads to the use of the cases of a second disease, different from the disease under investigation, as controls (Waller and Gotway 2004). In the spatial econometric field, a similar approach is applied in Arbia et al. (2012).

We use kernel smoothing to estimate the intensity $\lambda^*(u)$ of the controls, as implemented in the `density()` function of the **spatstat** package (Baddeley and Turner 2005) for the R computing environment (R Core Team 2014). Bandwidth is selected using cross-validation. Then, the intensity $\lambda(u)$ of the cases is estimated by a parametric model of the form

$$\lambda(u) = \lambda^*(u) \exp\left\{\alpha + \boldsymbol{z}(u)'\boldsymbol{\beta}\right\}, \tag{4}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m)'$ is the vector of regression coefficients.

Let Y_j be equal to 1 or 0 according to whether the jth household is a case or a control; under the assumption that both cases and controls are mutually independent IPPs, and conditional on their locations, we have

$$P(Y_j = 1|\boldsymbol{z}_j) = \frac{\exp(\alpha + \boldsymbol{z}_j'\boldsymbol{\beta})}{1 + \exp(\alpha + \boldsymbol{z}_j'\boldsymbol{\beta})}, \tag{5}$$

with $\boldsymbol{z}_j$ being the vector of spatial covariates on the location of the jth household. As equation (5) is a logistic regression model, its coefficients α and $\boldsymbol{\beta}$ can be estimated by maximum likelihood (Diggle et al. 2007).

We consider the following spatially referenced explanatory variables for model (4):

- *popdens*, the overall population density. Data used are the population counts for each of the 2,425 city's census tracts, as surveyed in the 2001 General Population Census (http://dawinci.istat.it/). Smooth population density has been estimated using Tobler's pycnophylatic interpolation algorithm (Tobler 1979).
- *rent*, minimum rent cost per square meter for private residential properties; these data are collected by the Italian Revenue Agency when rental contracts are registered. For the city of Catania, data are provided grouped within 28 homogenous geographical areas (Italian Revenue Agency - OMI database, http://www.agenziaentra te.gov.it/geopoi/omi/).

Table 2 reports the estimated regression coefficients, with their standard errors. Fig-
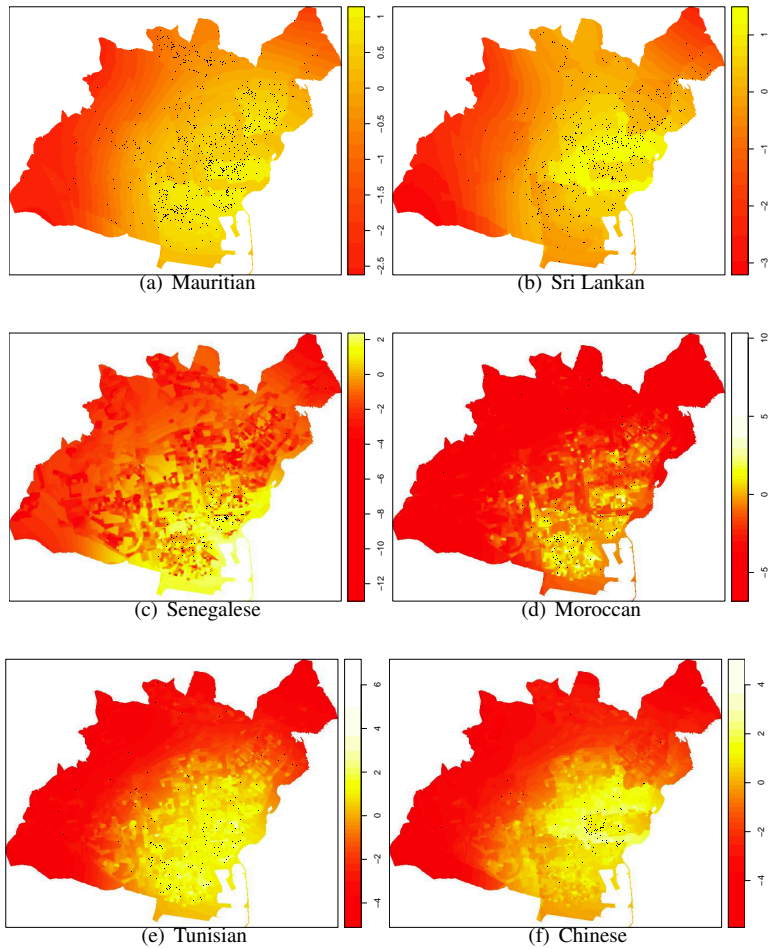
ure 2 shows the spatial distribution of $\ln(\widehat{\lambda}_i(u)/\widehat{\lambda}_i)$, where $\widehat{\lambda}_i$ is the average intensity of nationality i, $i = 1, \ldots, 6$; we use relative intensities to allow the comparison among the different nationalities. Furthermore, the log scale is used to make fractional changes equivalent.

Table 2: **Estimated regression coefficients and standard errors for $\ln(\hat{\lambda}_i(u))$.**

Nationality	Intercept	Rents	Popdens
Mauritian	0.64*** (0.17)	-0.21*** (0.04)	-0.02 (0.05)
SriLankan	-1.67*** (0.18)	0.31*** (0.04)	-0.01 (0.05)
Senegal	2.00*** (0.26)	-0.29*** (0.06)	-1.00*** (0.08)
Moroccan	-1.57*** (0.36)	-0.46*** (0.10)	0.75*** (0.09)
Tunisia	-2.45*** (0.32)	0.02 (0.08)	0.43*** (0.09)
Chinese	-4.25*** (0.34)	0.57*** (0.07)	0.29** (0.09)

All nationalities share a spatial trend, with higher values in the more central parts of the city (southernmost on the maps) and lower in the more peripheral areas. However, this spatial trend is much stronger for the vendors group than it is for the domestic workers group.

Figure 2: **Spatial distribution of** $\ln(\hat{\lambda}_i(u)/\hat{\lambda}_i)$

(a) Mauritian

(b) Sri Lankan

(c) Senegalese

(d) Moroccan

(e) Tunisian

(f) Chinese

4.2 Estimate of the spatial attraction

Having obtained the estimates $\widehat{\lambda}_i(u)$, $i = 1, \ldots, p$, for the $p = 6$ nationalities considered, we can now compute the estimates for the inhomogeneous K-function $\widehat{K}_{\text{inhom},i}(d)$.

Plots in Figure 3 show the behavior of $\widehat{K}_{\text{inhom},i}(d)$. Distances d are expressed in me-

ters. Plots also report the confidence envelopes referred to the null hypothesis of absence of interaction (see Section 2.4), at a significance level $\alpha = 0.05$ and based on $n = 999$ Monte Carlo simulations. As in Diggle et al. (2007), each simulation is obtained by randomly relabeling cases and controls, with spatially varying probabilities given by (5). Distances corresponding to the peaks of $\widehat{K}_{\text{inhom},i}(d)$ outside the envelopes are those at which there is significant spatial concentration or dispersion.

In Section 1 we mentioned the prominent role of social networks in attracting arriving migrants into specific neighborhoods and, also, channeling them into particular occupations (Gelderblom and Adams 2006).

MacDonald and MacDonald (1964) define the process in which "prospective migrants learn of opportunities, are provided with transportation, and have initial accommodation and employment arranged by means of primary social relationships with previous migrants" as chain migration. Chain migration subsidies the costs of migration; if the costs of migration are low, both economically and in terms of cultural distance, we may expect networks to play a less important role, and hence there is more individual migration. Empirical studies suggest that the willingness to help aspiring migrants is related to the lack of assimilation of the previous migrants into the new environment and to their desire of reestablishing, in the new town, social ties with co-villagers and relatives; furthermore, the plan of moving back to their country of origin in a short time can make migrants warier of refusing to help when approached (Banerjee 1983). So, it seems reasonable to expect that nationalities with stronger relational ties with the country of origin, and a more prominent intention to migrate back, may show a stronger chain migration and higher degrees of spatial attraction.

Mauritians show a significant spatial attraction at almost all the distances considered (Figure 3). Mauritian migration in Italy is concentrated in a few southern cities, of which Catania accounts for the highest portion, thus denoting a strong chain migration effect (King and Andall 1999). Mauritians perceive their stay in Italy as economically motivated and temporary, with an intention to migrate back to their home country as soon as they accumulated enough money (Crane 2004).

Sri Lankans show a significant spatial attraction at distances greater than 500 meters. Their migration took place with the beginning of the ethnic conflict between Sinhalese and Tamils, which started in 1983. Within Sicily, the two communities settled respectively in Catania and Palermo. A check in place of birth locations showed that the majority of Sri Lankans residing in Catania were born in the coastline region north of Colombo, in the cities of Marwila, Chilaw and Negombo, suggesting a relevant chain migration.

Figure 3: **Estimated inhomogeneous K-functions**

(a) Mauritian

(b) Sri Lankan

(c) Senegalese

(d) Moroccan

(e) Tunisian

(f) Chinese
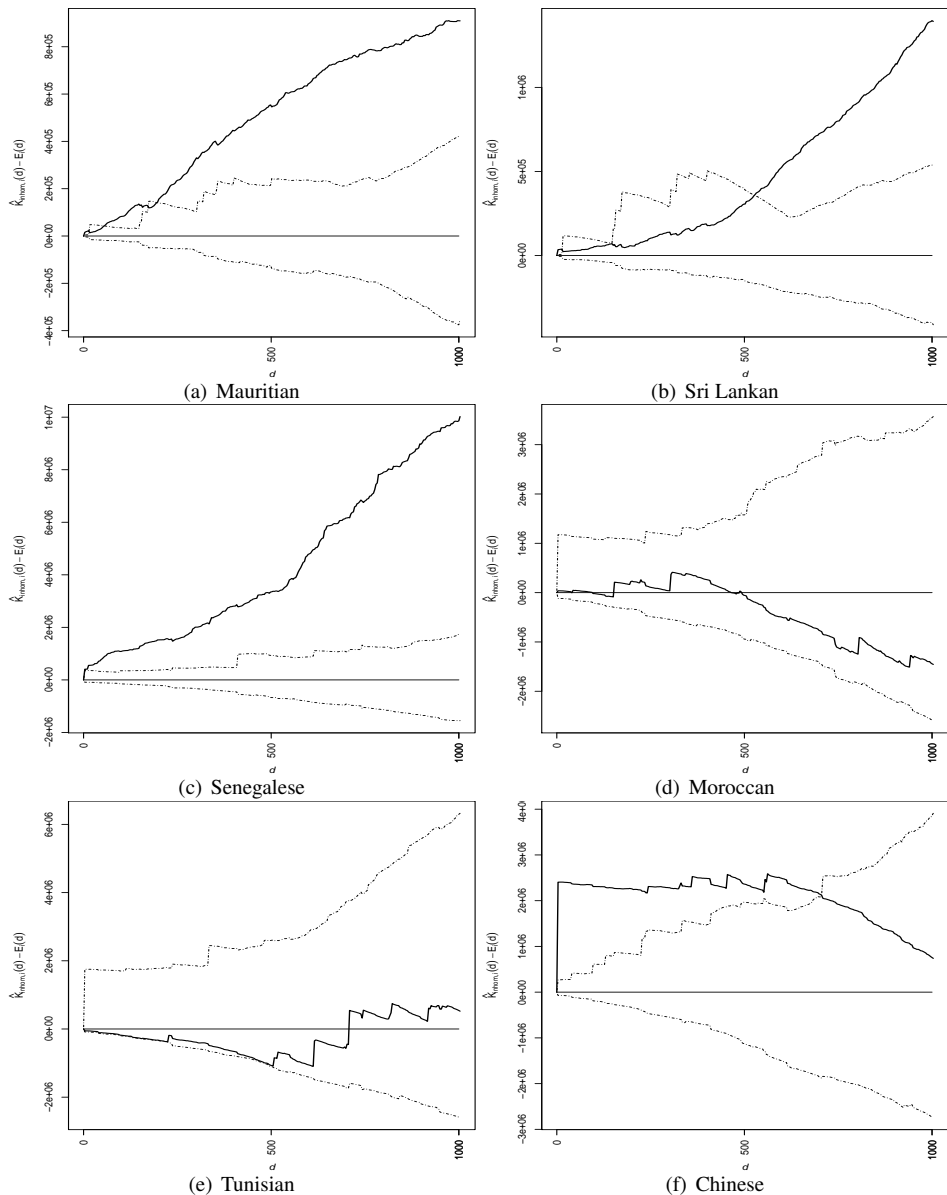
Note: Estimated inhomogeneous K-functions (solid black lines) and corresponding 95% confidence envelopes under the null hypothesis of absence of interaction, based on 999 Monte Carlo simulations. Values are centered with respect to the mean $E_i(d)$ of the simulations. Distances d are in meters.

Senegalese show a significant interaction at any distance. Most Senegalese migrants in Italy belong to the Mouride Sufi brotherhood, whose vertical and horizontal ties are reproduced in transnational networks, which help migrants to organize their business activities as well as their temporary settlement within the receiving contexts (Riccio 2008). Compared with other migrant communities, they show more mobility, following paths shaped by social, kinship and religious networks (Caritas Italiana, Fondazione Migrantes, and Caritas Diocesana di Roma 2004). Proofs of these mobility paths are on the Register Office database, which shows, over time, many Senegalese settling in and then leaving the same houses.

Tunisians show a settlement pattern consistent with random allocation. They are among the first immigrants who came to Sicily, with large communities settling in western Sicily in the second half of the sixties, due to geographical proximity and maritime linkages. Tunisians are widely distributed in Italy, with a variety of labor market roles in different parts of Italy: mostly in the primary sector (fishing, farming) and peddling in Sicily, in the service sector in central Italy and in factories in the north of Italy (King and Andall 1999). One reason for the absence of spatial attraction may be the low relevance of migratory chains in Tunisian migration to Sicily, which is reported to be mainly dictated by individual choice (Daly 2001, p. 193).

Moroccans' settlement pattern is also consistent with random allocation. As Tunisians, Moroccans also are a long-established group and they are widely dispersed in Italy, with different work specializations in different areas (King and Andall 1999). Moroccan migration to Italy began in the mid-1970s, but levels of migration remained marginal until the 1980s and grew significantly in the 1990s, when Moroccans became the largest foreign community in the country, before being overtaken by the Albanian and Romanian communities in recent years (Devitt 2013).

Chinese show a significant interaction at all distances less than $d = 750$ meters. The very high spatial attraction at lower distances might be explained with the fact that the most part of Chinese are settled in the buildings surrounding the open market area where they now run most of the stores; this area is starting being referred by the press as a local "Chinatown". Chang (2012, p. 181) reports Chinese migration in Italy to be a "calculated choice, made in the best interests of the family" and being characterized by "chain migration, ethnic enclave and familial duty". In addition to a great linguistic distance, Chinese live rather secluded (Mazza and Punzo 2016).

More in depth studies are needed to assess the reasons of the different settlement patterns produced by the different nationalities. However, our results seem consistent with a relevant impact of chain migration in spatial attraction among foreign migrants.

5. Conclusions

Traditional indices of spatial segregation rely on data aggregated by areal unit, typically census tracts (Mazza and Punzo 2015); see Matthews and Parker (2013) for a review of new spatially derived techniques to analyze segregation. These indices generally suffer from the modifiable area unit problem and do not allow us to distinguish between economically induced segregation, typically due to spatial inhomogeneity in the cost of living and job availability within different parts of the city, from spatial attraction among individuals of the same ethnicity. Schelling (1971) pinpoints the importance of this distinction, positing that even mild preferences for living with similar neighbors carry the potential to be strong determinants for residential segregation. In the case of minority migrants, spatial attraction may be stronger when the cost of settling is higher, either economically or in terms of cultural distance, and social networks may play an important role in channeling arriving migrants into specific neighborhoods and particular occupations.

We applied point-pattern analysis to analyze the settlement patterns of foreign migrants in the city of Catania (Italy); this approach was motivated by Almquist and Butts (2012) and Matthews and Parker (2013). We used the inhomogeneous K-function to assess spatial attraction among migrants while adjusting for spatial inhomogeneity.

To avoid the risk of confounding the two sources of clustering, spatial inhomogeneity is estimated following a case-control approach. We considered two groups, according to a work specialization that has developed among nationalities: domestic workers and vendors. The first group includes nationalities devoted to house work and care giving, the second group nationalities committed to peddling and retailing. We deemed it reasonable to assume that the environmental constraints within each occupational group are similar. We analyzed the six nationalities with the highest number of households.

Data at hand came from the population register, as recorded on August 30, 2005, and included individual residential addresses for all foreign residents. For each of the six nationalities analyzed, we considered the set of all the other nationalities of the same occupational group as a control group. Spatial inhomogeneity has been estimated for each nationality by means of a parametric model of the spatial intensity of the controls and two explanatory variables related to the overall population density and to the minimum rent cost per square meter. The spatial intensities of the controls has been estimated nonparametrically. All nationalities share a spatial trend, with higher values in the more central parts of the city and in the central business district, and lower in the more peripheral areas. However, this spatial trend is much stronger for the vendors group than it is for the domestic workers group.

Finally, we computed the inhomogeneous K-function, with confidence envelopes based on 999 Monte Carlo simulations referred to the null hypothesis of absence of interaction ($\alpha = 0.05$). Each simulation was obtained by randomly relabeling cases and controls. A significant spatial attraction has been found for Sri Lankan, Mauritians, Sene-

galese, and Chinese. Conversely, the settlement patterns of Tunisians and Moroccans are consistent with random allocation. More in depth studies are needed to assess the reasons of the different settlement patterns produced by the different nationalities. However, these results seem consistent with a relevant impact of chain migration in spatial attraction among foreign migrants.

References

Almquist, Z.W. and Butts, C.T. (2012). Point process models for household distributions within small areal units. *Demographic Research* S13(22): 593–632. doi:10.4054/DemRes.2012.26.22.

Altavilla, A.M. and Mazza, A. (2012). On the analysis of immigrant settlement patterns using quadrat counts: The case of the city of Catania (Italy). *Advances and Applications in Statistics* 29(2): 111–123.

Arbia, G., Espa, G., Giuliani, D., and Mazzitelli, A. (2012). Clusters of firms in an inhomogeneous space: The high-tech industries in Milan. *Economic Modelling* 29(1): 3–11. doi:10.1016/j.econmod.2011.01.012.

Baddeley, A. (2010). *Analysing spatial point patterns in R*. Tech. rep., CSIRO. www. csiro. au/resources/pf16h.html.

Baddeley, A.J., Møller, J., and Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica* 54(3): 329–350. doi:10.1111/1467-9574.00144.

Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial point patterns: Methodology and applications with R*. London: Chapman and Hall/CRC Press. http://www.crcpress.com/Spatial-Point-Patterns-Methodology-and-Applications-with-R/Baddeley-Rubak-Turner/9781482210200/.

Baddeley, A. and Turner, R. (2000). Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics* 42: 283–322. doi:10.1111/1467-9574.00144.

Baddeley, A. and Turner, R. (2005). Spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software* 12(6): 1–42. http://www.jstatsoft.org/v12/i06/.

Banerjee, B. (1983). Social networks in the migration process: Empirical evidence on chain migration in India. *The Journal of Developing Areas* 17(2): 185–196.

Bartlett, M. (1963). The spectral analysis of point processes. *Journal of the Royal Statistical Society. Series B* 25(2): 264–296.

Busetta, A., Mazza, A., and Stranges, M. (2016). Residential segregation of foreigners: an analysis of the Italian city of Palermo. *Genus* 71(2–3): 177–198.

Caritas Italiana, Fondazione Migrantes, and Caritas Diocesana di Roma (2004). *Immigrazione, Dossier Statistico 2004. XIV Rapporto*. Rome: Idos.

Chang, A. (2012). 20th century Chinese migration to Italy: The Chinese diaspora presence within European international migration. *History Compass* 10(2): 179–190.

doi:10.1111/j.1478-0542.2011.00833.x.

Clark, W.A. and Fossett, M. (2008). Understanding the social context of the Schelling segregation model. *Proceedings of the National Academy of Sciences* 105(11): 4109–4114.

Crane, K. (2004). *Governing migration: Immigrant groups' strategies in three Italian cities – Rome, Naples and Bari*. Tech. rep., Psychoanalytic Institute for Social Research, Rome.

Daly, F. (2001). The double passage. In: King, R. (ed.). *The Mediterranean passage: Migration and new cultural encounters in Southern Europe*. Liverpool: Liverpool University Press.

Devitt, C. (2013). Circular economic migration between Italy and Morocco. In: Triandafyllidou, A. (ed.). *Circular migration between Europe and its neighbourhood: Choice or necessity?* Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199674510.003.0005.

Diggle, P.J., Gómez-Rubio, V., Brown, P.E., Chetwynd, A.G., and Gooding, S. (2007). Second-order analysis of inhomogeneous spatial point processes using case-control data. *Biometrics* 63(2): 550–557. doi:10.1111/j.1541-0420.2006.00683.x.

Freeman, L., Pilger, J., and Alexander, W. (1971). *A measure of segregation based upon spatial arrangements*. University of Hawaii.

Gelderblom, D. and Adams, W. (2006). The limits and possibilities of migrant networks. In: Kok, P., Gelderblom, D., Oucho, J.O., and Van Zyl, J. (eds.). *Migration in South and Southern Africa: Dynamics and determinants*. Cape Town: HSRC Press: 227–248.

Harvey, D. (1966). Geographic processes and the analysis of point patterns: Testing models of diffusion by quadrat sampling. *Transactions of the Institute of British Geographers XL* 81–95. doi:10.2307/621570.

Heinrich, L. (1991). Goodness-of-fit tests for the second moment function of a stationary multidimensional Poisson process. *Statistics* 22(2): 245–268. doi:10.1080/02331889108802308.

King, R. and Andall, J. (1999). The geography and economic sociology of recent immigration to Italy. *Modern Italy* 4(2): 135–158. doi:10.1080/13532949908454826.

Kumar, N. (2007). Spatial sampling design for a demographic and health survey. *Population Research and Policy Review* 26(5-6): 581–599. doi:10.1007/s11113-007-9044-7.

Kumar, N., White, M.J., and Singh, G. (2013). Multi-group individual level measures of segregation. *SSRN*. http://ssrn.com/abstract=2205714.

Licciardello, O. and Damigella, D. (2011). Immigrants and natives: Ways of constructing new neighbourhoods in Catania, Sicily. In: Armbruster, H. and Meinhof, U.H. (eds.). *Negotiating multicultural Europe: Borders, networks, neighbourhoods*. London: Palgrave Macmillan: 141–158. doi:10.1057/9780230346475$_7$.

MacDonald, J.S. and MacDonald, L.D. (1964). Chain migration, ethnic neighborhood formation and social networks. *The Milbank Memorial Fund Quarterly* 42(1): 82–97. doi:10.2307/3348581.

Matthews, S. and Parker, D.M. (2013). Progress in spatial demography. *Demographic Research* 28(10): 271–312.

Mazza, A. and Punzo, A. (2015). On the upward bias of the dissimilarity index and its corrections. *Sociological Methods and Research* 44(1): 80–107.

Mazza, A. and Punzo, A. (2016). Dealing with omitted answers in a survey on social integration of immigrants in Italy. *Mathematical Population Studies.* Forthcoming.

Møller, J. and Waagepetersen, R.P. (2003). *Statistical inference and simulation for spatial point processes*. London: CRC Press.

Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich: Geobooks.

R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.

Reardon, S.F. (2006). A conceptual framework for measuring segregation and its association with population outcomes. In: Oakes, J. and Kaufman, J. (eds.). *Methods in social epidemiology*. San Francisco: Jossey-Bass: 169–192.

Riccio, B. (2008). West African transnationalisms compared: Ghanaians and Senegalese in Italy. *Journal of Ethnic and Migration Studies* 34(2): 217–234. doi:10.1080/13691830701823913.

Ripley, B.D. (1981). *Spatial statistics*. New York: John Wiley & Sons. doi:10.1002/0471725218.

Ripley, B.D. (1991). *Statistical inference for spatial processes*. Cambridge: Cambridge University Press.

Schelling, T.C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology* 1(2): 143–186. doi:10.1080/0022250X.1971.9989794.

Tobler, W.R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association* 74(367): 519–530. doi:10.1080/01621459.1979.10481647.

Waller, L.A. and Gotway, C.A. (2004). *Applied spatial statistics for public health data*,

vol. 368. New York: John Wiley & Sons. doi:10.1002/0471662682.

Wong, D. (2009). The modifiable areal unit problem (MAUP). In: Fotheringham, A. and Rogerson, P. (eds.). *The SAGE handbook of spatial analysis*. London: SAGE Publications: 105–123. doi:10.4135/9780857020130.n7.