MEASURING THE SPATIAL CONCENTRATION OF THE MAIN FOREIGN COMMUNITIES RESIDING IN ITALY USING A NEW APPROACH

Massimo Mucciardi, Giovanni Pirrotta, Mary Ellen Toffle

Abstract. In the field of population studies the spatial concentration of population reflects many issues. Surely one of the most important is the residential segregation of the foreign populations which is closely linked to the settlement pattern of the population. Starting from previous research on spatial concentration proposed by Mucciardi and Benassi (2023) and subsequently by Benassi *et al.* (2023a), the present paper develops new insights into spatial concentration applied to the main foreign communities in Italy. The aims of the contribution are as follows: (i) to detect the level of spatial concentration of the main foreign communities residing in Italy in 2003, 2011 and 2021; (ii) to evaluate the role of country of citizenship in shaping such levels and dynamics. To achieve aims (i) and (ii) we applied a spatial version of the Gini index, called the Spatial Gini Index (SGI) and the relative curve called Spatial Lorenz Curve (SLC). The results show that the level of the spatial concentration for each foreign group remains almost stable over time (with rare exceptions). In contrast, the level of the spatial concentration for the main foreign communities is lower than the Italian population.

1. Introduction

As stated in a recent paper (Benassi *et al.*, 2023b), spatial concentration is an important aspect when analysing various distributions and patterns pertaining to immigrant communities. According to the important study conducted by Massey and Denton (1988), one of the most indicative elements of residential segregation is that of the actual concentration of the immigrant population. Furthermore, to a recent study conducted by Mucciardi and Benassi (2023), the concentration of the population in terms of space continues to be a strong indicator of urbanization. The phenomenon of immigration in Italy has demonstrated a rapid growth cycle in the last years. For example, there was a very small population of foreigners residing in Italy and most of them had illegal or irregular status (Strozza, 2004). After regularization procedures were put into force, the numbers of foreigners residing in

Italy increased rapidly to 1.89 million individuals in the first part of 2003 (Blangiardo, 2005). But migration intensified between 2003 and 2021: foreign residents increased from 4.32 million in 2011 to 5.03 million in 2021 with a net immigration of about 3.14 million people between 2003 and 2021 (ISTAT, 2023). In addition, there was a change in the countries of origin. East-west groups coming from the Middle East started replacing the south-north immigration pattern of low-income migrants from Asia and Africa (Benassi *et al.*, 2023b). In Italy foreigners are concentrated in certain areas of the country, especially in the large metropolitan cities. However, not all groups of foreigners tend to concentrate or settle in the same way: it is evident that the territorial distribution of the different nationalities is connected not only with the duration of stay (communities of recent immigration vs. communities of previous immigration) and with the stability of the communities on the territories but also with the different migration models. So, the foreign groups follow different settlement patterns and show different levels of spatial concentration. In the field of population studies, the spatial concentration of population reflects many other issues. Surely one of the most important is the residential segregation of the foreign populations which is closely linked to the settlement pattern of the population. Starting from previous research on spatial concentration proposed by Mucciardi and Benassi (2023) and subsequently by Benassi *et al.* (2023a), the present paper developed a new insight of spatial concentration applied to the main foreign communities in Italy. The aims of the contribution are the following: (i) to detect the level of spatial concentration of the main foreign communities residing in Italy in 2003, 2011 and 2021; (ii) to evaluate the role of country of citizenship in shaping such levels and dynamics. To achieve aims (i) and (ii) we applied a spatial version of the Gini index, called the Spatial Gini Index (SGI) proposed by Mucciardi and Benassi (2023) and subsequently by Benassi *et al.* (2023a). SGI is derived from the Lorenz curve (Lorenz, 1908), and it is based on comparing how the contribution in terms of “connectivity” and “variability” varies as the geographical distance between spatial units increases. The paper is organized as follows. In the next section we show a short review of the new procedure while in section 3 we present the results obtained. Some final remarks conclude the paper.

2. A new approach to measure spatial concentration: a short review

Certainly, in the study of concentration the best known and most widespread index is the Gini index (G). As already pointed out in Benassi *et al*. (2023b), despite the large dissemination of the index G, it has also been criticized. When applied to phenomena involving the study of variables in a territorial context, essentially this index does not take into consideration the geographical nature of the phenomenon, being fundamentally an "aspatial" index (Reardon and O'Sullivan, 2004). To overcome this limitation some alternative measures and approaches have been continually proposed by scholars (Arbia and Piras 2009; Dawkins 2004; Rey and Smith 2013; Panzera and Postiglione 2020, Türk and Östh, 2023).

Here we briefly describe the method proposed by Mucciardi and Benassi (2023) and subsequently by Benassi *et al.* (2023a).

This approach is based on comparing how the contribution in terms of "*connectivity*" ($J_{(k)}$) and "*variability*" ($V_{(k)}$) varies as the geographical distance k between spatial units increases: if the variable observed is not dependent on space, the variations between the connectivity and variability components should not differ much from each other. The idea is to consider buffer or threshold distances (k) capable of progressively creating partitions of the territory (or territorial subsets). These partitions identify neighbouring and non-neighbouring units such that each partition is disjoint from the others and the sum of all the elements of all the partitions coincides with the number of all the possible pairs between the n spatial units (Mucciardi and Benassi, 2023; Benassi *et al.*, 2023). To satisfy these conditions we use the MaxMin distance method (Mucciardi, 2008). The properties of the territorial partitions of the MaxMin method makes the procedure compatible with the graphical representation of the Gini index according to the Lorenz curve approach (Lorenz, 1908). We called this index the "Spatial Gini Index" (SGI) and the relative curve "Spatial Lorenz Curve" (SLC).

Basically we can have three scenarios (Figure 1):

A) when the cumulative variability contribution in terms of variability $V_{(k)}$ is be larger than the cumulative connectivity contribution in terms of connectivity $J_{(k)}$ as the distance k increases, then 0≤SGI <0.5 (case of negative spatial autocorrelation, Figure 1A)[1].

B) when the cumulative variability contribution $V_{(k)}$ is equal to the cumulative connectivity contribution $J_{(k)}$ as the distance k increases, then SGI is perfectly equal to 0.5 (case of spatial no autocorrelation, Figure 1B)[2];

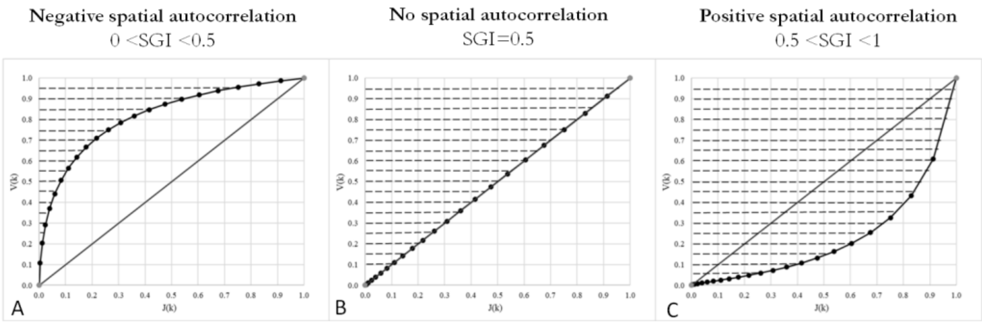
C) when the cumulative variability contribution in terms of variability $V_{(k)}$ is smaller than the cumulative connectivity contribution in terms of connectivity $J_{(k)}$ as the distance k increases, then 0.5<SGI ≤1 (case of positive spatial autocorrelation, Figure 1C).

[1] For more details on the calculation of the SGI and on the construction of the SLC curve, see Mucciardi and Benassi (2023) and Benassi *et al*. (2023).

[2] We can think this case when in the Lorenz curve there is exactly equidistribution of a variable (for example income).

It is important to highlight the link between *spatial concentration* and *spatial autocorrelation*. Previous studies (Mucciardi and Benassi, 2023; Benassi *et al.*, 2023a) show that when a variable (e.g. human population) is uniformly distributed in the entire territory, the SGI tends to 0.5 and the SLC tends to the configuration of Figure 1B (case of no spatial autocorrelation). If the variable is not distributed uniformly throughout the entire territory, two sub-cases may arise: i) in the case of non-uniform distribution (or in any case with high variability) in the contiguous areas, SGI tends towards values lower than 0.5 and the SLC tends to the configuration of Figure 1A (case of negative spatial autocorrelation); ii) in the case of uniform distribution (or in any case with low variability) in the contiguous areas, SGI tends towards values greater than 0.5 and the SLC tends to the configuration of Figure 1C (case of positive spatial autocorrelation).

Figure 1 – *Three scenarios for SGI and SLC (dashed line): 1) case of negative spatial correlation – 0<SGI <0.5 (A); 2) case of no spatial correlation – SGI=0.5 (B); 3) case of positive spatial correlation 0.5<SGI <1 (C). The diagonal (solid line) represents the theoretical condition of spatial equality.*



Source: Benassi et al., 2023a

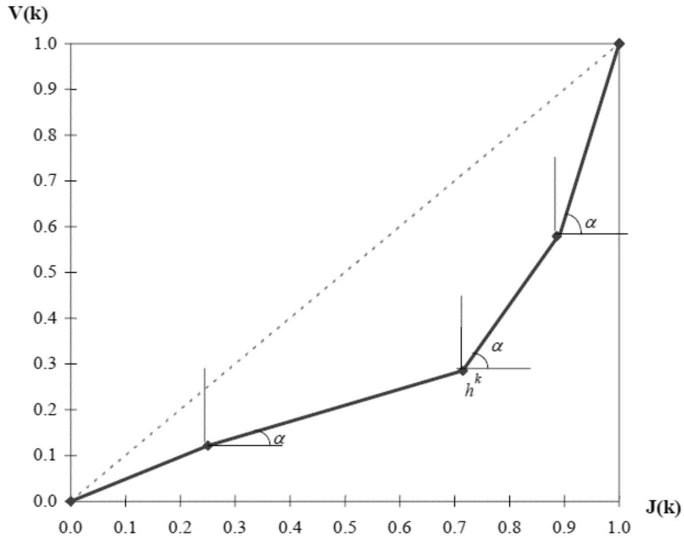
Another interesting feature of this new approach is the possibility of measuring the level of spatial autocorrelation at each distance (k) using the arctan function (AF). In fact, from a geometric point of view, these three forms of spatial autocorrelation may be assessed, as the distance k varies, by considering the tangent of the angle formed by the straight line with the *x*-axis (Mucciardi and Benassi, 2023; Benassi *et al.*, 2023a). In formula:

$$\tan^{k}(\alpha) = \frac{V_{(k)}}{J_{(k)}} \quad k = 1 \ldots t. \tag{1}$$

where t coincides with the maximum territorial distance (beyond this distance the MaxMin distance method stops).

We can have 3 cases (Figure 2): a) $\tan^{\mathrm{k}}(\alpha) < 1$ ($\alpha < 45°$) indicating positive spatial autocorrelation; b) $\tan^{\mathrm{k}}(\alpha) = 1$ ($\alpha = 45°$) indicating no spatial autocorrelation; c) $\tan^{\mathrm{k}}(\alpha) > 1$ ($\alpha > 45°$) indicating negative spatial autocorrelation. To convert the tangent function into terms of angle (degrees) we use the AF function[3].

Figure 2 – Values assumed by the angle α with varying values of distance (k) (dashed line represents case of no spatial autocorrelation - $tg^{k}(\alpha) = 1$ ($\alpha = 45°$).



Source: Mucciardi and Benassi (2023)

Furthermore, two indicators are derived from SGI for comparing the value of the spatial concentration of a variable x with respect to the value of the spatial concentration of a second variable y_{ref} usually set as a reference. We called these indicators "Delta SGI" (ΔSGI_x) and "SGI Rate" ($RSGI_x$).

We can write the following formula:

$$\Delta SGI_x = SGI_x - SGI_{ref} \text{ and } RSGI_x = \frac{SGI_x}{SGI_{ref}} \quad (2)$$

with $-1 \leq \Delta SGI_x < 1$ and $0 \leq RSGI_x < \infty$ ($SGI_{ref} \neq 0$)

[3] We recall that the AF is the inverse of the tangent function and it returns the angle whose tangent is a given number.

Considering that SGI is a dimensionless index, the ΔSGI_x can be considered as an absolute variation of the spatial concentration of a variable x with respect to the spatial concentration of the variable taken as a reference. Instead $RSGI_x$ can be considered as a relative variation of the spatial concentration of a variable x with respect to the spatial concentration of the variable taken as a reference. Clearly, the more the spatial concentration of the variable x tends to the spatial concentration of the variable set as a reference, the values of ΔSGI_x and $RSGI_x$ tend to 0 and 1 respectively. Therefore, in the following paragraph we will use this property to verify the similarity/dissimilarity of a foreign community compared to the Italian one in terms of spatial concentration.

3. Measuring the spatial concentration of the main foreign communities: principal results

3.1 Data and MaxMin algorithm output

Before analyzing the results obtained, it is necessary to examine the data used. Data is provided by the Italian National Institute of Statistics (ISTAT, 2023) and it refers to the population usually residing in the Italian municipalities broken down by country of citizenship. Territorial boundaries of 7903 municipalities have been reconstructed so they remain stable over time assuring a correct comparison. Regarding the results of the MaxMin algorithm applied to Italian municipalities, we obtain a total of 163 (k) buffer distances ranging from the minimum value of 16.3 km ($k = 1$) to the maximum value of 1223.9 km ($k \equiv t = 163$)[4].

3.2 Principal results

We calculate the SGI for the first 10 largest foreign communities in 2021 by years: 2003, 2011 and 2021 (Table 1). According to formula (2), two indicators are derived from SGI to compare the value of the spatial concentration of the single foreign community (SGI_{FOR}) with the spatial concentration of the Italian communities (SGI_{ITA}). So, for the 10 foreign communities considered we have:

$$\Delta SGI_i = SGI_{ITA} - SGI_{FOR_i} \text{ and } RSGI_i = \frac{SGI_{FOR_i}}{SGI_{ITA}} \qquad \text{with } i = 1 \dots 10 \qquad (3)$$

[4] All elaborations are based on a new ad hoc library developed and implemented in Python. The library can be downloaded at the following link: https://github.com/gpirrotta/spatial-gini-index (Benassi *et al.*, 2023a).

Table 1 summarizes the three spatial concentration indicators for the 10 foreign communities. As we can see (Table 1 and Figure 3) the level of the spatial concentration for each foreign collectivity (SGI_03, SGI_11 and SGI_21) remains almost stable over time with rare exceptions (Bangladesh and Pakistan). In contrast the level of the spatial concentration for the main foreign communities is lower than the Italian population. The value of SGI for Italians varies from 0.485 to 0.483 in the three years considered, substantially showing (as expected) no spatial concentration and stability over time (population tending towards territorial homogeneousness). Instead, Pakistanis, Egyptians and Chinese communities show lower levels of SGI and consequently negative spatial autocorrelation probably due to concentrations of these populations in the North of the Italian territory (Table 1 and Figure 3).

Table 1 – *SGI, $\Delta\boldsymbol{SGI_i}$ and $\boldsymbol{RSGI_i}$ for Romania (ROM), Morocco (MOR), Albania (ALB), China (CHI), Ukraine (UKR), India (IND), Philippines (PHI), Bangladesh (BAN), Egypt (EGY), Pakistan (PAK) citizenship – Years 2003 (03), 2011(11) and 2021 (21). Italy (ITA) is reported for reference.*

Citizen	SGI_ 03	SGI_ 11	SGI_ 21	ΔSGI_ 03	ΔSGI_ 11	ΔSGI_ 21	RSGI_ 03	RSGI_ 11	RSGI_ 21
ROM	0.465	0.466	0.465	0.019	0.019	0.018	0.960	0.960	0.963
ALB	0.444	0.437	0.441	0.041	0.049	0.042	0.915	0.900	0.913
MOR	0.454	0.458	0.457	0.031	0.027	0.026	0.937	0.944	0.947
CHI	0.415	0.420	0.413	0.070	0.065	0.070	0.855	0.866	0.854
UKR	0.479	0.467	0.464	0.006	0.018	0.019	0.987	0.962	0.961
IND	0.451	0.451	0.466	0.034	0.035	0.017	0.930	0.928	0.965
PHI	0.434	0.437	0.440	0.051	0.048	0.043	0.895	0.900	0.910
BAN	0.486	0.469	0.467	-0.002	0.016	0.016	1.003	0.967	0.968
EGY	0.417	0.417	0.417	0.068	0.069	0.066	0.863	0.858	0.864
PAK	0.412	0.412	0.434	0.073	0.073	0.049	0.849	0.849	0.898
ITA	*0.485*	*0.485*	*0.483*	-	-	-	-	-	-

Source: Our elaboration on ISTAT data

Moreover, if these foreign communities are compared with the Italian population, we obtain the highest ΔSGI_i and lowest $RSGI_i$ values in all the years considered in this analysis (Table 1 and Figure 4). The other foreign communities seem to have values of the three indices tending towards the values of the Italian population, especially the Ukrainian and Romanian communities (see the values of ΔSGI_i and $RSGI_i$ tending to 0 and 1 respectively). Now let's examine the SLC and the AF. We recall that SGI is a global index while through the SLC and AF it is possible to inspect the spatial autocorrelation values for each distance k.

Figure 3 – *Comparison of the value of SGI by citizenship (Italians are plotted for reference).*

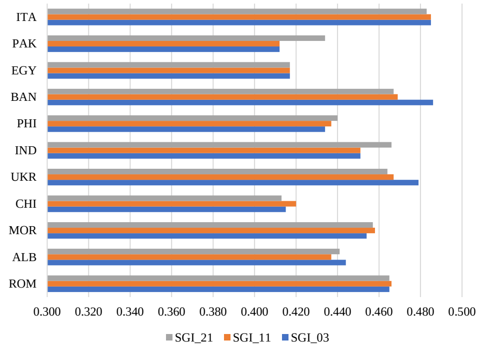
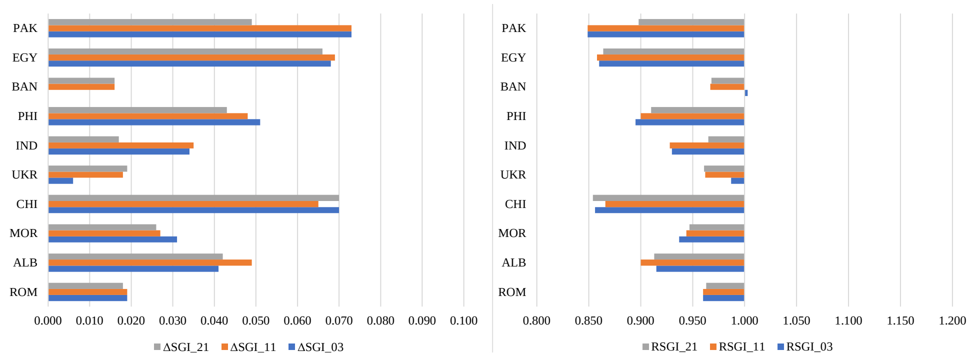


Figure 4 – *Comparison of the value of ΔSGI_i and $RSGI_i$ by citizenship (Italians are the reference).*



If we consider SLC (Figure 5) and the value of the AF (Figure 6) the differences appear greater[5]. For instance, comparing the SLC and the AF of Romanian and Egyptian (year 2021) we notice completely different trends that confirm, albeit in a different way, the settlement models known in the literature (Benassi *et al.*, 2020; Ferrara *et al.*, 2010): a dispersed model (Romanian) and a concentrated model (Egyptian). However, even in terms of SLC, the communities of Bangladesh and Pakistan do not remain stable over time (date not shown). To check the accuracy of the SGI, we compare the SGI results with the ones obtained using global concentration indices known in the literature: Delta index (DEL) (Hoover, 1941; Duncan *et al.*, 1961) and Absolute concentration index (ACO) (Massey and Denton,

[5] For reasons of space, we have reported only the main ones. All SLC can be downloaded at the following link: Supplementary files (SIEDS 2023).

1988). Although the results might appear controversial, we find a significant negative correlation between SGI and DEL (-0.58) and ACO (-0.48) applying the latter to the same database. In our opinion, the explanation for this result lies in the fact that both DEL and ACO are basically aspatial indices like the Gini index. Greater population concentration implies low territorial uniformity of the population and consequently low SGI values.

Figure 5 – *Comparison between SLC (from left to right): CHI vs ITA; EGY vs ITA; PHI vs ITA; MOR vs ITA; UKR vs ITA and ROM vs ITA (Years=2021).*

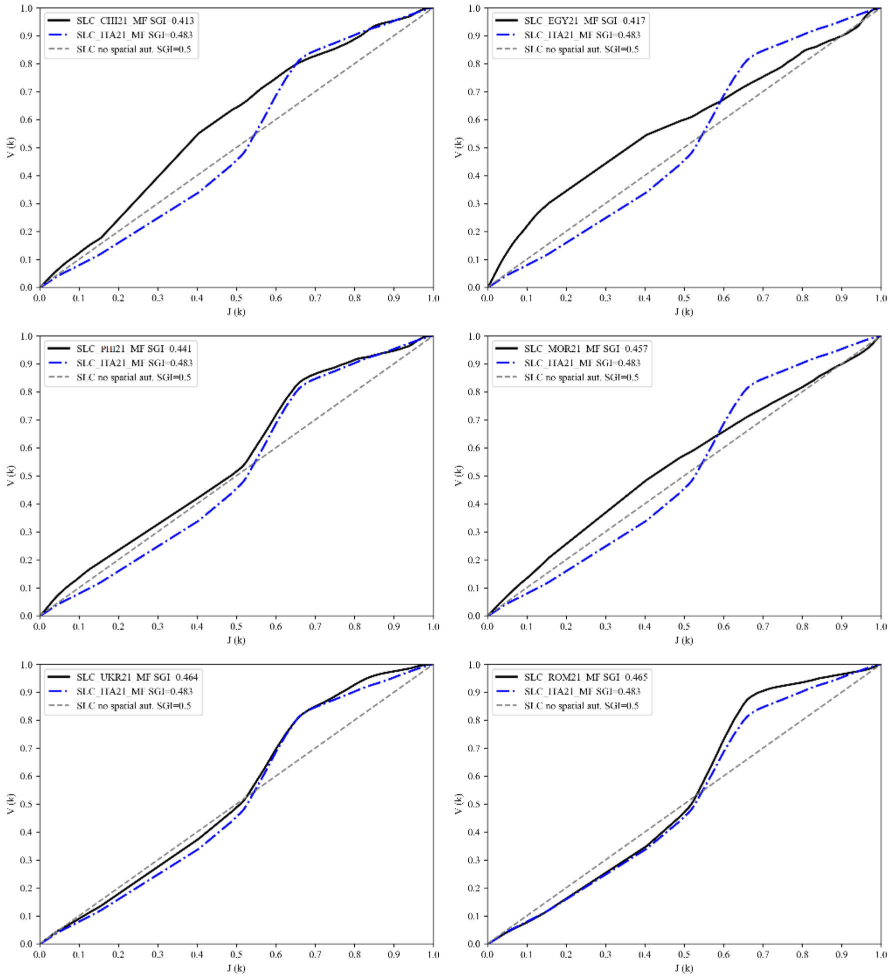
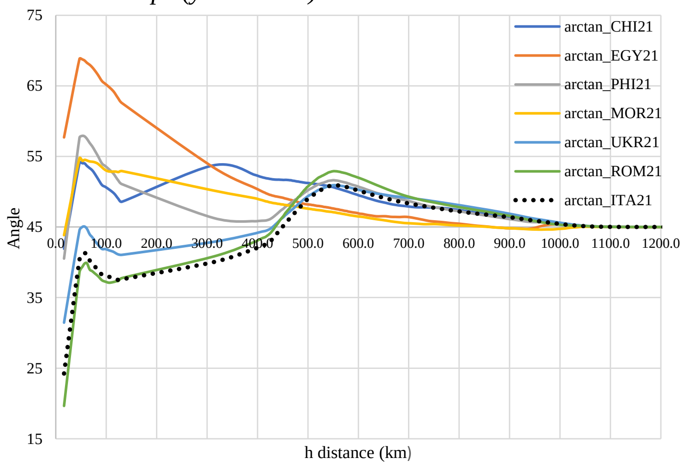


Figure 6 – *Value of the angle (degree) for the AF for the CHI, EGY, FIL, MOR, UKR and ROM citizenships (years 2021).*



4. Final remarks

In accordance with a new measure of spatial concentration proposed by Mucciardi and Benassi (2023) and Benassi *et al.* (2023a), the paper attempts to analyze the settlement models of the main foreign communities residing in Italy in the periods 2001, 2011 and 2021. The main results obtained can be summarized as follows. Compared to the traditional methodology in the field of population studies, SGI, SLC and AF represent an "*integrated tool*" for measuring spatial segregation through a global measure of spatial concentration (SGI), its graphical representation makes it possible to evaluate how a population is distributed in the territory (SLC) and the level of spatial autocorrelation of the population (AF). The application of the procedure to Italian municipal data makes it possible to differentiate the settlement models between foreign communities. Each community, even if the indicators of spatial concentration indicate stability over time (with some rare exceptions), seems to have its own settlement model: the populations of Romanians and Ukrainians demonstrate a settlement model similar to that of the Italians whereas the Egyptian and Chinese settlement models are different. To conclude, population statistics in general and especially spatial concentration measures are topics of study that can be useful in many ways. EU and national integration studies may use data obtained from spatial concentration measurement to understand the present condition of the ongoing efforts to integrate foreign populations residing in Italy. Knowing where national groups congregate also helps to understand the demand for various occupations including construction, agriculture, domestic services and industrial labour markets. This in turn can assist in urban planning and the creation of education, employment, and social programs. It is hoped that the research methodology and results presented here will be useful to other EU member states in the effort to promote integration and social cohesion. From methodological point of view, other improvements in the model are planned by the authors. We will implement a neighbourhood system based on other buffer distances or spatial weight matrices. Furthermore, the extension of inferential methods to test the significance of SGI (i.e. Monte Carlo test) is being planned.

References

ARBIA G., PIRAS G. 2009. A new class of spatial concentration measures. *Computational Statistics and Data Analysis*, Vol. 53, No. 21, pp. 4471-4481. https://doi.org/10.1016/j.csda.2009.07.003.

BENASSI F., MUCCIARDI M., PIRROTTA G. 2023a. Looking for a new approach to measuring the spatial concentration of the human population. *Journal of Official Statistics*, forthcoming.

BENASSI F., MUCCIARDI M., CARELLA M., NACCARATO A., SALVATI L. 2023b. Spatial segregation in action? An empirical assessment of population concentration of foreigners and nationals in Italy, 2002-2018. *International Migration Review*. https://doi.org/10.1177/01979183231170808.

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*, Vol. 101, https://doi.org/10.1016/j.habitatint.2020.102200.

BLANGIARDO G. C. 2005. I processi di immigrazione: dall’illegalità alla regolarizzazione’. In: Livi Bacci M. (Ed.), *L’incidenza economica dell’immigrazione*, Quaderni Cesifin, n. 20, Giappichelli.

DAWKINS C. J. 2004. *Measuring the spatial pattern of residential segregation*. Urban Studies, Vol. 414, pp. 833–851. https://doi.org/10.1080/0042098042000194.

FERRARA R., FORCELLATI L., STROZZA S. 2010, Modelli insediativi degli immigrati stranieri in Italia, *Bollettino della Società Geografica Italiana*, Vol. 13, No. 3, pp. 619-639.

DUNCAN O. D., CUZZORT R. P., DUNCAN B. 1961. *Statistical geography: Problems in analyzing areal data*. New York. The Free Press of Glencoe.

HOOVER E. 1941. Interstate redistribution of population, 1850-1940. *Journal of Economic History*, Vol. 1, pp. 199-205.https://doi.org/10.1017/S0022050700052980.

ISTAT 2023. Retrieved from: http://dati.istat.it/Index.aspx?QueryId=19125.

LORENZ M. O. 1905. Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, Vol. 970, pp. 209-219. https://doi.org/10.2307/2276207.

MASSEY D. S., DENTON N. A. 1988. The dimensions of residential segregation, *Social Forces*, Vol. 67, No 2, pp. 281-315.

MUCCIARDI M., BENASSI F. 2023. Measuring the spatial concentration of population: a new approach based on the graphical representation of the Gini index. *Quality & Quantity,* Vol. 57, pp.5193-5211. https://doi.org/10.1007/s11135-022-01607-2.

MUCCIARDI M. 2008. Use of a flexible weight matrix in a local spatial statistic. In: Giusti A., Ritter G. and Vichi M (Eds.), *Classification and Data Mining,* Vol. 5. Naples: Ed. Sc. Italiane, pp. 385-388.

PANZERA D., POSTIGLIONE P. 2020. Measuring the spatial dimension of regional inequality: An approach based on the Gini correlation measure. *Social Indicators Research*, Vol. 1482, pp. 379-394. https://doi.org/10.1007/s11205-019-02208-7.

REARDON S. F., O'SULLIVAN D. 2004. Measures of spatial segregation. *Sociological Methodology*, Vol. 341, pp. 121-162. https://doi.org/10.1111/j.0081-1750.2004.00150.x.

REY S. J., SMITH R. J. 2013. A spatial decomposition of the Gini coefficient. *Letters in Spatial and Resource Sciences*, Vol. 62, pp. 55-70. https://doi.org/10.1007/s12076-012-0086-z.

STROZZA S. 2004. Estimates of the Illegal Foreigners in Italy: A Review of the Literature, *International Migration Review*, Vol. 38, No. 1, pp. 309-331.

TÜRK U., ÖSTH J. 2023. Introducing a spatially explicit Gini measure for spatial segregation. Journal of Geographical Systems, Vol.25, pp.469-488. https://doi.org/10.1007/s10109-023-00412-1.

Massimo MUCCIARDI, Department of Cognitive Science, Education and Cultural Studies, University of Messina, Italy, massimo.mucciardi@unime.it.

Giovanni PIRROTTA, IT Staff, University of Messina, Italy, gpirrotta@unime.it.

Mary Ellen TOFFLE, Department of Political and Juridical Sciences, University of Messina, Italy, maryellen.toffle@unime.it.