9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

Social-economic segregation dynamics at neighborhoods scale

Mobility, GPS, Income Segregation, Social Mixing, Urban Planning

Extended Abstract

Income segregation is commonly defined as the uneven interaction of communities inside the urban environment. Cities face income segregation as a threat to economic growth and social stability, as poverty and crime rates increase [1]. Income segregation has also a significant impact on a person's access to opportunities and services [2]. Its effects drive not only static spatial residential separation [3] but also a different commute duration and insufficient social mixing. These lead to worse quality of life for lower-income and fewer opportunities for productive contact between people with different incomes.

Segregation, urban topology, and human mobility are interwoven: their dynamics affect daily activities, and citizens' interaction with the places of the cities [4, 5]. People's whereabouts depend on necessities: we commute for work, shop grocery, or leisure, according to our lifestyle [6]. We are driven by Points Of Interest (POIs), at a given time, in a given neighborhood; they affect our social activity and track the level of segregation. As people move within the city, our experience of social mixing will change during the day: do neighborhoods change inclusiveness in relation to their topological characteristics? What are the topological features that promote inclusiveness?

The challenge at hand is to associate income segregation with topological features of the neighborhood: the POIs, as the urban dimension's services, and the people's mixing, related to the time of the day. We analyze income segregation focusing on the dynamics of social mixing in *time and space* in the city of Milan for eight months. We propose space-time metrics to fully characterize segregation. From the spatial effect, we define a representative score of Attractivity, Liveability, and Accessibility (ALA-score). Accessibility is represented by Velocity Score, the ease of getting to a given place [7]. For attractivity we define: the Fitness [8] of a neighborhood, that represents both the sparsity and uniqueness of POIs categories, the diversity of categories, the diversity of prices, the median of prices, and reviews. Finally, three quantities that measure liveability: the number of supermarkets and their diversity, and the number of schools. These ALA metrics cluster the neighborhoods into three groups indicating their goodness (Fig. 1).

Along the mobility dimension, we grasp the effect of social mixing by focusing on the percentage of income in a given neighborhood in a given time. This dynamic is designed by leveraging on trajectory Location-Based Services (LBS) [9] from 21 thousand users in a period of ten months coupled with a geolocated dataset of rent for square meters, used as a proxy of user's income. We are not seeing any representativeness bias in the real population living in a neighborhood, according to census data.

The neighborhoods' spatio-temporal dynamics of the social mixing permit to capture time and ALA-score dependent social interaction changes. We show that ALA metrics of a given neighborhood interact differently with social classes in a time-dependent way (Fig. 2). There is high spatial segregation at night, due to residential segregation, but more social mixing, especially pairwise, during the day when residential segregation is relaxed. Neighborhoods, where interaction with the middle class is present, are more inclusive: medium income will attend places in common both with low and high income. Segregation is still present depending on the cluster being considered: it will depend on the type of facilities a neighborhood offers, used at different times by different incomes.

We define the temporal mixing of a neighborhood as the temporal pattern of a function of the Gini coefficient. We cluster the temporal mixing and we get three clusters, one inclusive, one segregated, and one mostly inclusive but with segregated hours, such as weekend afternoons (Fig. 3 (**B**)). We finally study the similarity between the clusters given by the ALA-score and those of temporal mixing: not all the neighborhoods of the ALA-score clusters end up in the same mixing cluster (Fig. 3 (**C**)). To understand the neighborhood's features that drive the inclusion, we used a null model that randomizes income and preserves users' whereabouts. In this way through the distributions, Fig. 3 (**D**), we gauge how statistically significant ALA-metrics are to describe inclusivity.

A more accurate picture of segregation is provided: we associated it with simple topological metrics, such as Points Of Interest price or public transport accessibility. By working on these metrics, and implementing some modifications on neighborhood levels, we contribute to urban theories to make the city more inclusive.

References

[1] Kawachi, I., Kennedy, B. P., Lochner, K. and Prothrow-Stith, D. *Social capital, income inequality, and mortality.*, Am. J. Public Health 87, 1491 (1997).

[2] Tammaru T, Marcińczak S, Aunap R, et al. *Inequalities and segregation across the long-term economic cycle: An analysis of south and north European cities.*, IZA Discussion Paper, 10980 (2017)

[3] Massey, Douglas S. and Denton, Nancy A., *The Dimensions of Residential Segregation*, Social Forces **67**, (1988).

[4] Moro, E. and Calacci, D. and Dong, X. et al., *Mobility patterns are associated with experienced income segregation in large US cities.*, Nature Communications **12**, (2021).

[5] Tóth, G. and Wachs, J. and Di Clemente, R. et al. , *Inequality is rising where social network segregation interacts with urban topology*, Nature Communications, **12**, (2021).

[6] R. Di Clemente, M. Luengo-Oroz, M. Travizano, S. Xu, B. Vaitla and M. C. González *Sequences of purchases in credit card data reveal lifestyles in urban populations*, Nature Communications, **9** (2018)

[7] Indaco Biazzo, Bernardo Monechi and Vittorio Loreto , *General scores for accessibility and inequality measures in urban areas*, The Royal Society, **6**, (2019).

[8] Andrea Tacchella, Matthieu Cristelli, Guido Caldarelli, Andrea Gabrielli, Luciano Pietronero, *A new metrics for countries' fitness and products' complexity*, Scientific reports, **2**, (2012).

[9] Jochen S, Agnès V. *Location-Based Services.* , Elsevier Inc, **24**, (2004)

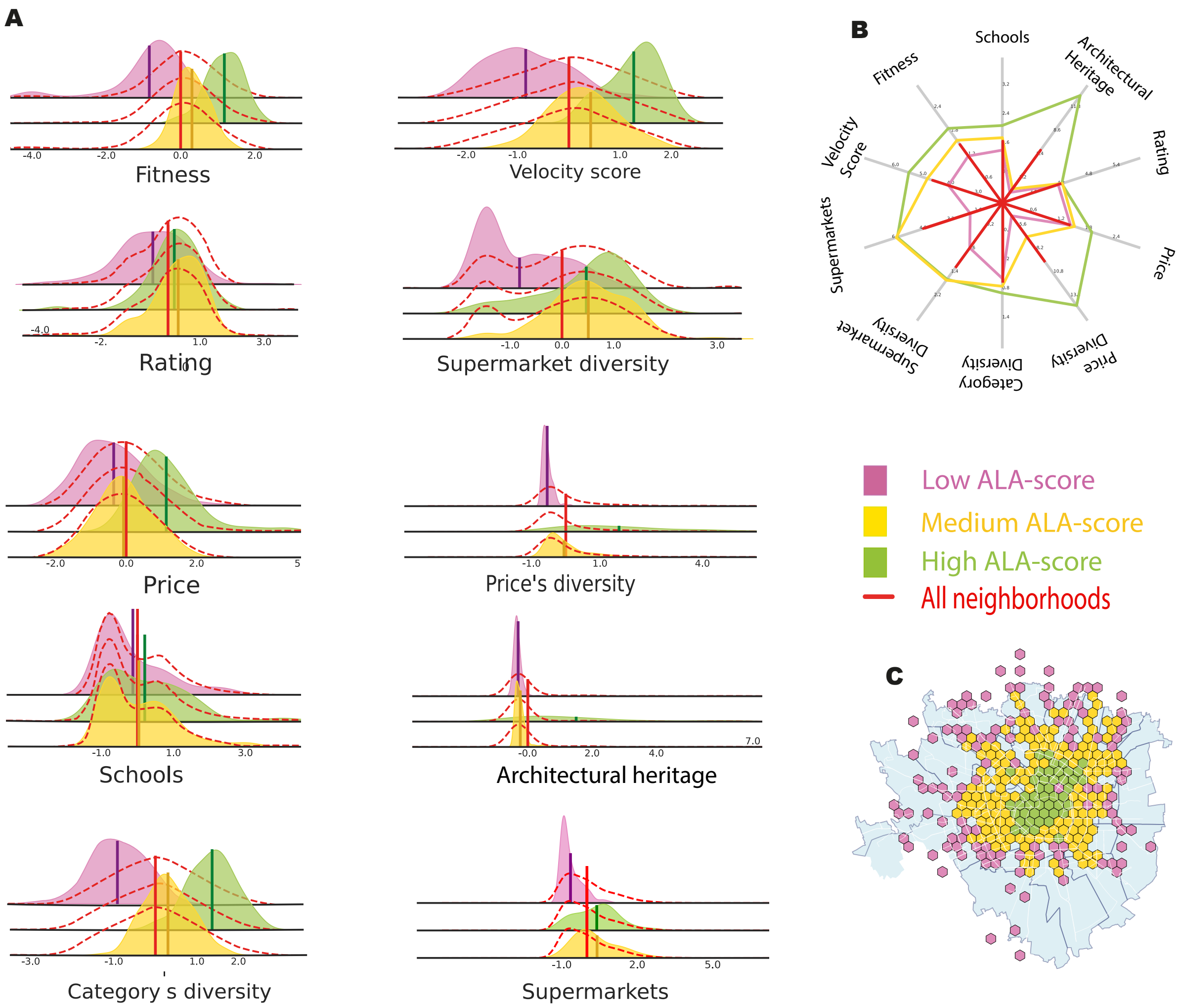

Figure 1: **ALA-score of neighborhood**. The hexagons represent the neighborhoods . (**A**) Distribution of ALA metrics in zscore among neighborhoods: fitness, velocity score, category diversity, median price, price diversity, rating, supermarket diversity, number of supermarkets, schools, and architectural heritage. The three clusters are represented by filled colors, the lines of the respective color indicate the median and the dotted red line stands for the distribution of all neighborhoods. (**B**) Comparison of the median of the ALA metrics within each cluster with respect to the median of all neighborhoods in red. (**C**) Location of clusters in the city of Milan.

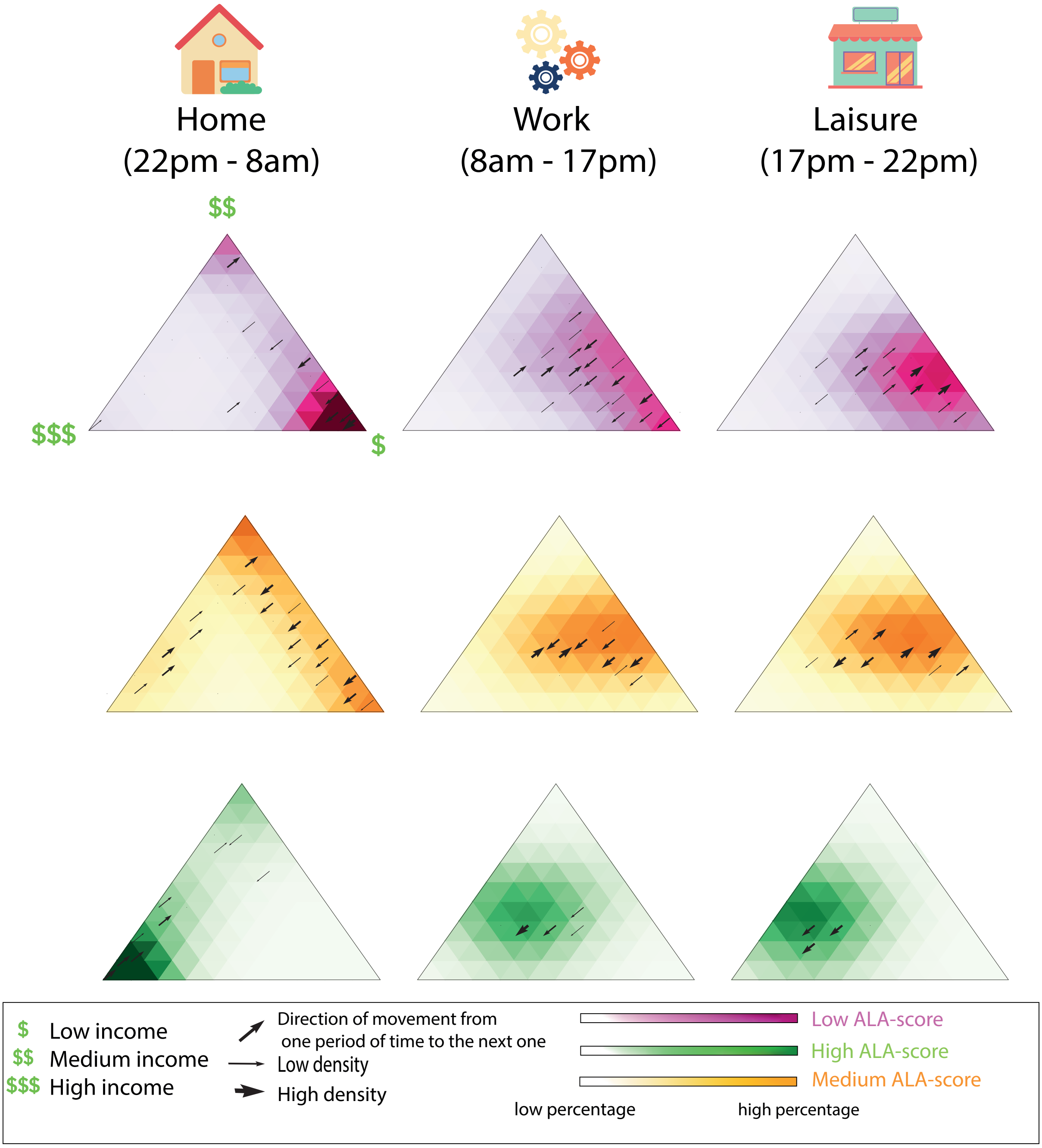

Figure 2: **Movements of people in ALA-score clusters.** At the vertices of the triangles there are incomes, the more a neighborhood is visited by an income the closer it is to one of the vertices; the more the neighborhood is well distributed among the incomes the closer it is to the center of the triangle. People may be in a neighborhood because they live there, because they work there, or for other reasons. Depending on the time of the day, the neighborhood will have a different percentage of income. Each row indicates the cluster to which the neighborhood belongs, and each column indicates the time of day being observed (home at night, work during the day, other in the evening). The arrows indicate the movement that will occur in the following time. The thinness of the arrows indicates the density of the neighborhood in which the same phenomenon happens.

.

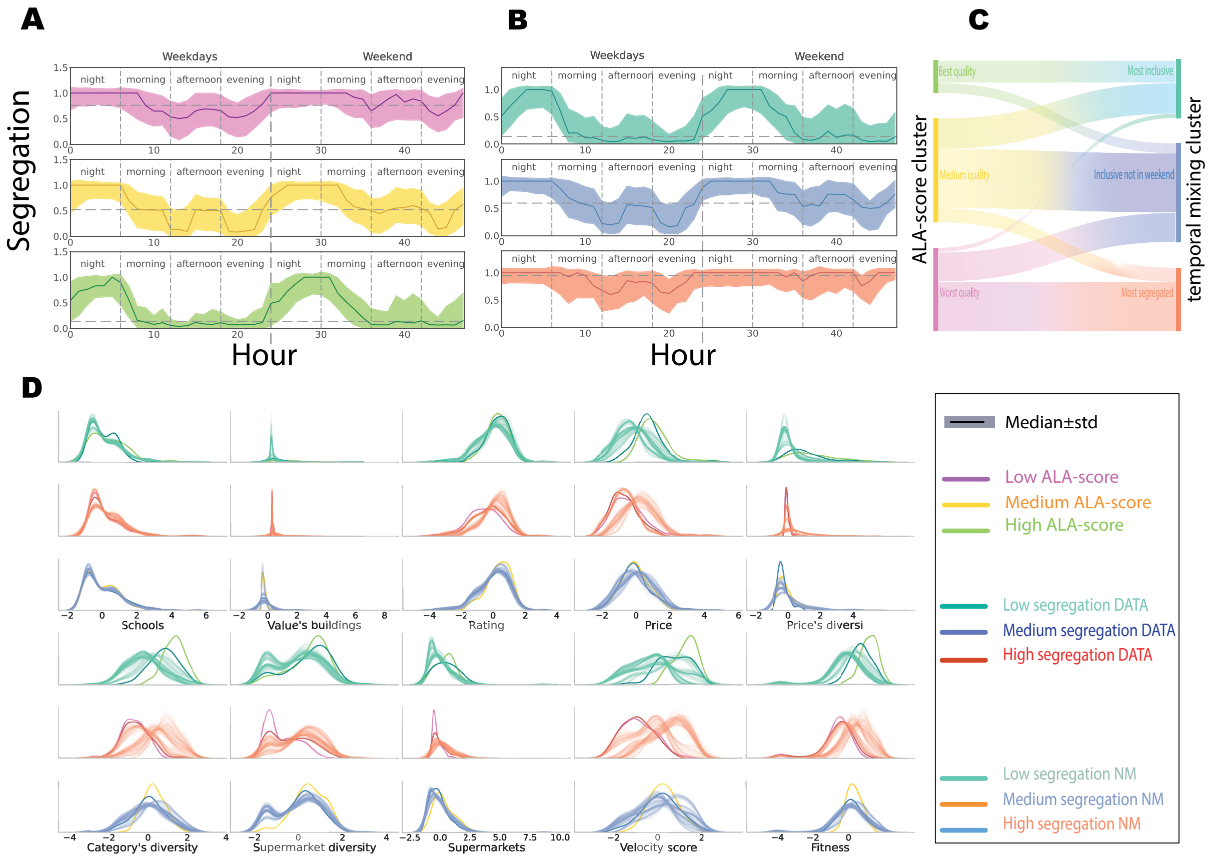

Figure 3: **Importance of ALA features in income segregation.** (**A**) Segregation profile for ALA cluster during the weekdays and the weekend. The solid line is the median and the shaded area is the standard deviation. (**B**) Segregation profile for the new cluster, the temporal mixing cluster, during the weekdays and the weekend. The solid line is the median and the shaded area is the standard deviation. (**C**) Partitioning of the neighborhood between ALA-score clusters and temporal mixing clusters. (**D**) After randomizing the incomes we recalculated the temporal mixing and assigned the cluster each time. In this plot, there are the contributions of each ALA feature for each cluster, ALA-score cluster, REAL temporal mixing cluster, and NULL MODEL (NM) temporal mixing cluster.

.