Computers, Environment and Urban Systems 86 (2021) 101585

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Are we breaking bubbles as we move? Using a large sample to explore the relationship between urban mobility and segregation

Souneil Park [a,*], Taylor M. Oshan [c], Abdallah El Ali [b], Alessandro Finamore [a]

[a] Telefonica Research, Barcelona, Spain
[b] Centrum Wiskunde & Informatica (CWI), Amsterdam, Netherlands
[c] University of Maryland, College Park, MD, United States

ARTICLE INFO

Keywords:
Mobility segregation
Spatial interaction models
Urban mobility
Urban informatics

ABSTRACT

Segregation often dismantles common activity spaces and isolates people of different backgrounds, leading to irreconcilable inequalities that disfavour the poor and minorities and intensifies societal fragmentation. Therefore, segregation has become an increasing concern and topic of research with studies typically concentrating on the residential communities of a particular racial or socioeconomic group. This paper enhances the residential view of segregation and examines the topic in the context of urban mobility. Specifically, it expands upon prior research by employing large-sample, seamless telecommunication logs of London, UK to provide a holistic view of mobility across the entire socioeconomic spectrum. A method is developed to transform the data to flows between geographic areas with different socioeconomic statuses. Spatial interaction models are then calibrated to examine the impact of both geographical distance and socioeconomic distance on the deterrence of flows and the analysis is extended to analyze the interaction of the two factors. Overall, socioeconomic distance is found to have a subtle effect compared to geographical distance. However, different effects are observed depending on the socioeconomic distance between flows and the deterrence of mobility tends to be the greatest when both physical and socioeconomic distance are high, suggesting that both factors may play a role creating and maintaining segregation.

1. Introduction

Segregation reinforces rising inequality through long-term choices such as where to live, work, or attend school, as well as daily mobility choices, such as places to shop, socialize, or hospitals to visit. Segregation can further permeate our lives as market mechanisms more deeply penetrate society (Sandel, 2012). For example, it may take root gradually, but communities may quickly become accustom to rising levels of isolation. It is inherently complicated to understand segregation as it is the outcome of many factors that depend upon spatial and demographic context. As a result, the issue of segregation is typically framed in terms of a particular racial group, socioeconomic class or geography, which may only partially characterize the 'bubbles' that encapsulate lived experiences.

A large body of work on segregation is facilitated by the availability of traditional surveys (e.g., the census) of residential areas, schools, and workplaces of the population where the geodemographics are relatively static. Conspicuous patterns often emerge, though these efforts do not capture the experience of highly mobile urban populations that are dynamic across space and time. (Charles, 2003). There have been a number of attempts to capture mobility-based segregation using social media (Wang, Edward, Phillips, & Sampson, 2018), public transportation logs (Lathia, Quercia, & Crowcroft, 2012) and call detail records (CDRs) of mobile phones (Silm & Ahas, 2014). These studies have confirmed the existence of segregation (Wang et al., 2018), demonstrated that it may vary over time (Silm & Ahas, 2014)), and suggested that denser samples of urban mobility across the population are needed to paint an accurate picture of mobility-based segregation.

This paper builds on previous efforts and overcomes some limitations by approaching mobility-based segregation through the use of a large sample of the urban population across the entire socioeconomic spectrum. Specifically, it takes advantage of seamless city-scale mobility traces extracted from a European mobile network operator for the city of London, UK. This allows for continuous location approximation as antennas record the pairing events with devices (e.g., network attachment, detachment, and handovers). It is also combines the mobility traces with

* Corresponding author at: Torre Telefónica Diagonal 00 Plaza de Ernest Lluch i Martín, 5 08019 Barcelona, Spain.
E-mail address: souneil.park@telefonica.com (S. Park).

https://doi.org/10.1016/j.compenvurbsys.2020.101585
Received 16 April 2020; Received in revised form 3 November 2020; Accepted 5 December 2020
Available online 2 January 2021

census data, enabling a comprehensive segmentation of the population by socioeconomic status. The density and diversity of our sample is a major advantage over previous research that focused on particular racial/ethnic groups or selected a sub-set of classes (e.g., rich vs. poor). Consequently, it is possible to present a more holistic characterization of mobility-based segregation.

The remainder of this paper is organized as follows: some related work is first reviewed before introducing the data and methods employed in this research; then the results are described; next, a discussion of the results and their significance in relation to mobility-based segregation is provided; and finally, some concluding remarks and potential extensions of this work are put forth.

2. Background and previous work

Segregation is a central concern for contemporary political, economical, and cultural well-being. It does not simply denote socioeconomic, racial, or ethnic differences within an urban space; rather it is an underlying structure of societal fragmentation and irreconcilable inequality (United Nations Centre for Human Settlements, 2001). One research approach has been to understand the proliferation of segregation as a consequence of growing income inequality. For example, one study examined census data from 1970 to 2000 for 100 metropolitan areas in the U.S. and demonstrated that income inequality contributes towards income segregation, which is defined as the "uneven geographic distribution of income groups within a certain area" (Reardon & Bischoff, 2011). Recent research also provides evidence on the evolution of segregation: for example, many studies observed an expansion of segregation (Musterd, Marcińczak, van Ham, & Tammaru, 2017; van Ham, Marcińczak, Tammaru, & Musterd, 2015; van Kempen & Murie, 2009) whereas one study observed a decrease in racial segregation in the U.S. (Lee, 2016).

There is a growing concern that segregation reinforces an unfair playing-field in terms of access to social capital, which ultimately disfavours the poor and ethnic minorities (Wilson, 1990), and various efforts are being made to understand societal and individual outcomes of segregation and how to address the associated problems. The Moving to Opportunity (MTO) experiment (Sanbonmatsu, Ludwig, Katz, et al., 2011) utilized a representative sample where 4600 low-income families were recruited and provided with housing vouchers for relocation. The experiment resulted in a number of major findings such as a link between movements and better health outcomes. A follow-up study further examined the long-term neighborhood effects and found that moving to a lower-poverty neighborhood during adolescence was associated with increased college attendance and reduced single parenthood rates (Chetty, Hendren, & Katz, 2016). In addition, other research efforts reported evidence that segregation is associated with inequality in health care (Bach, Pham, Schrag, Tate, & Hargraves, 2004; Smith, Feng, Fennell, Zinn, & Mor, 2007), education and employment outcomes (Cutler & Glaeser, 1997), and safety and crime rates (Krivo & Peterson, 1996). Consequently, understanding segregation, its drivers, and its outcomes remains an important topic across disciplinary perspectives.

In particular, segregation is often investigated through a geographical lens that seeks to measure a variety of socioeconomic characteristics across cities, metropolitan areas, or countries. This line of inquiry tends to vary according to a diversity of factors, such as how the population is segmented (e.g., black-white (Massey, Rothwell, & Domina, 2009)), the geographical area of interest, scale of spatial units, or specific manifestations of segregation (e.g., racial segregation at work (Ellis, Wright, & Parks, 2004)). However, a common thread across many geographical studies of segregation is that they focus on a static population, such as the distribution of residencies (Charles, 2003; Farrell, 2008; Katz & Lang, 2004) or work places (Ellis et al., 2004). The typical approach is to take a geographical division and/or a distribution of socioeconomic categories and to employ a measure or index that quantifies the magnitude of segregation between the categories across the geographical divisions. Census data is often employed because it offers detailed socioeconomic and racial attributes across large study areas and provides clear geographic divisions (Charles, 2003; Massey et al., 2009; Reardon & Bischoff, 2011). Furthermore, many indices have been developed and used to quantify and compare the degree of segregation. Massy and Denton (Massey & Denton, 1988) conducted an extensive review that surveyed and categorized 20 segregation indices. For instance, a measure of category "evenness" compares the ratio of majority/minority groups in a spatial unit, with the measure becoming the highest when no members of the minority and majority groups are located in the same spatial unit. Since these methods are designed to measure the overlap of the group distributions, an important task is how the groups are defined.

A more recent perspective recognizes that residential and work locations are only a portion of where social interaction occurs (van Ham & Tammaru, 2016). In contrast to the traditional geographical approach, mobility-based segregation expands the focus to include individuals' full activity space – the exhaustive set of places an individual might experience on a day-to-day basis. Daily activity spaces provide a basis for understanding actual exposure of individuals to different social groups. Systematic differences in activity spaces between social groups implies varying day-to-day territories, decreasing the likelihood of interaction. In contrast, overlapping activity spaces imply longer interaction amongst individuals with potentially diverse backgrounds (Park & Kwan, 2018). Furthermore, there is a growing body of research that recognizes the importance of mobility in defining the spatial context of observations. For example, Kwan (Kwan, 2013) discusses the importance of incorporating time and mobility into geographic research to move beyond the static spatial perspective. Specifically, segregation is pointed out as a topic that can be refined when it is viewed through the lens of mobility and activity spaces.

A number of recent studies explored segregation across social spheres and routinised activities (Netto, Soares, & Roberto, 2015) or daily lived experiences of different social groups (Zhang & Wang, 2019), which has practical implications for bridging interactions across population groups and for the equitable production of social capital (Wilson, 2012). In particular, overlapping activity spaces and co-presence in urban environments facilitate interactions, enabling individuals to experience each other (Netto et al., 2015). While the type of interactions can range from exclusion and tolerance to sharing physical space and "welcoming of the other" (Levinas, 1979; Netto et al., 2015), co-presence opens the potential to create mutual understanding of lived experiences (Legeby, 2013).

Investigating mobility not only exposes additional dimensions of segregation, but it also raises new questions about relationships with the urban environment, such as the distribution of economic opportunities (Kain, 2004) or public transportation networks (Lucas, 2012). Though understanding connections to these factors calls for further research, the focus here is directly on a large-scale analysis of mobility-based segregation and the dynamic nature of mobility in contrast to assuming a static residential population. In addition, an exploratory analysis of the mobility-based segregation is necessary before examining additional factors.

A major challenge to exploring mobility-based segregation is capturing the mobility of individuals across segments of the population, for which various data sources have been employed. For instance, CDRs are commonly used as they enable large-scale analyses by tracking the approximate location of where mobile devices connect to antennas (Dannamann, Sotomayor-Gómez, & Samaniego, 2018; Silm & Ahas, 2014). For example, through CDRs Silm and Ahas (Silm & Ahas, 2014) observe a temporal variation in the degree of segregation in the capitol city of Estonia. They find variations along the diurnal cycle, and also between weekdays and weekends, arguing that the degree of segregation is lower when considering mobility data than when considering census data. CDRs are also used to predict socioeconomic indicators of individuals (Pappalardo, Pedreschi, Smoreda, & Giannotti, 2015) and geographic areas (Smith-Clarke, Mashhadi, & Capra, 2014), though these studies do not directly focus on segregation. One drawback of CDRs is that they are inherently limited in their ability to accurately identify trip destinations of daily mobility due to the sparsity of records across different time spans. This issue is further compounded by the increasing usage of messaging apps instead of SMS, which are not recorded in CDRs.

Another source of data for studying mobility-based segregation at a large-scale is geo-tagged social media. Wang et al. (Wang et al., 2018) studied neighborhood isolation in 50 American cities through geocoded Twitter data of 400,000 users generated over 18 months. Segmenting the users into poor vs. non-poor and white, black, and Hispanic ethnicity, the work analyzes the diversity and range of trips, and inter-group exposure. Although the data source allows GPS-level spatial accuracy, the temporal sparsity described for CDRs is similarly problematic here. In addition, social media sources are likely prone to demographic bias as active users are not typically representative of the general population (Jiang, Li, & Ye, 2019). Moreover, individuals' social media usage habits could vary widely and may not be representative of routine mobility activities (Lindqvist, Cranshaw, Wiese, Hong, & Zimmerman, 2011).

Public transportation logs have also been used to explore the segregation of mobility in London (Lathia et al., 2012). By aggregating trips between stations, Lathia et al. measure homophily based on the socioeconomic attributes of where the stations are located and find that well-off areas are visited by more diverse populations whereas people from well-off areas tend not to visit deprived areas. However, since the study is restricted to a particular mode of transportation, it is noted that homophily cannot be assumed to exist more generally.

These data sources previously employed to study mobility-based segregation are limited in their ability to consistently capture fine-grained mobility activity across space and time, represent the general population, or both. Consequently, this research builds on previous efforts to explore mobility-based segregation by taking advantage of the high-frequency mobility traces of nearly 2 million individuals sampled continuously across space and time for a single city. Östh et al.'s recent work (Östh, Shuttleworth, & Niedomysl, 2018) uses a similar type of telecommunication data for measuring mobility-based segregation in several Swedish metropolitan areas and finds that mobility tends to reduce segregation, particularly in central areas. However, only data from a single day (24 h) and a binary classification (i.e., poverty vs. wealth) were used to estimate exposure. Therefore, the data source employed here enables a more comprehensive estimation of population mobility and socioeconomic classification and is further described in the following section.

3. Data

3.1. Study area: London, UK

The city of London, UK was selected as the study area for this research for two main reasons. First, London provides an interesting study area because the city hosts a large and diverse population that is contained within a highly connected metropolitan area, facilitating daily urban interactions and offering the opportunity to make rich observations with respect to mobility-based segregation. The second reason is more practical and is based on the availability of both high resolution mobility data and detailed geodemographic data. These data sources are described in more detail below.

3.2. Mobility traces

In order to obtain high frequency mobility data, this research takes advantage of the network access logs of a major UK telecom operator, which keeps track of the antenna that a mobile device is connected to. The dataset logs all access events of mobile devices to the cellular network, which occur whenever a device establishes or updates a channel with the network. For example, when the device attaches to or detaches from the network, or the device is handed over between antennas. The mobile device frequently interacts with the network through such activities in order to maintain high quality service. While multiple antenna technologies (i.e., 2G/3G/4G) are available, in general, modern phones are designed to automatically join the fastest data connection available, making them prioritize 4G over 2G or 3G unless a user explicitly configures the phone not to use 4G. In addition, 4G provides better data and voice services by an order of magnitude, and its deployment is much denser than older technology. This typically results in a 4G (over 2G or 3G) connection with an antenna that is physically proximal to the mobile device for greater bandwidth and smaller last-mile latency.

The high coverage and the density of the antennas permit the location of mobile devices to be identified at a fine-grained geography (i.e., Lower Super Output Areas), which is further described in the subsequent section. In contrast to the CDRs or geo-social media check-ins more used in prior research, this dataset also provides continuous mobility traces with a high temporal granularity. In particular, a subset of network activity was employed that covers the Greater London area during the month of January 2018.

To ensure privacy, every user is assigned a unique random ID and no personal information, such as name, age, or address, is included in the network event data. Moreover, the main analysis of this research is conducted using only aggregated data, further decreasing privacy risks.

3.3. Index of multiple deprivation

The neighborhoods of London can be classified socioeconomically based on the *Index of Multiple Deprivation*[1] (IMD) survey conducted by the UK government. Rather than using only on a single attribute, the IMD consists of seven distinct dimensions: 1) income deprivation, 2) employment deprivation, 3) health deprivation and disability, 4) education, skills and training deprivation, 5) crime, 6) barriers to housing and services, and 7) living environment deprivation. As a result, the IMD is employed frequently in many domains (e.g., (Mcgillion, Pine, Herbert, & Danielle, 2017; Rivas, Kumar, & Hagen-Zanker, 2017)) as an inclusive metric for socioeconomic status and is the socioeconomic variable used in this research. Furthermore, the IMD is available for a fine-grained spatial division of England, called Lower-layer Super Output Areas (LSOAs). LSOAs are small areas designed to have a similar number of socially homogeneous inhabitants. An LSOA has an average of approximately 1500 residents or 650 households. There are 32,844 LSOAs in England and 4834 LSOAs in London.

Based on the IMD, a ranking amongst the 32,844 LSOAs across England is created and grouped into 10 equal-sized deciles that divide the units from the most deprived to the least deprived. For example, the LSOAs in decile 1 fall within the most deprived 10% nationally while the LSOAs in decile 10 fall within the least deprived 10%. Fig. 1 illustrates the distribution of London LSOAs across the national IMD deciles. It can be seen that the number of LSOAs in the extreme deciles (i.e., 1 and 10) are significantly lower than other deciles and that a majority of LSOAs belong to the lower-to-middle deciles (i.e., 2, 3, and 4). Meanwhile, the spatial distribution of the IMD deciles across the LSOAs of London is heterogeneous and displays clusters of various sizes for both high and low values (Fig. 2). Since both the IMD data and mobile network event data can be obtained for LSOAs, it is possible to process and combine the two data sources to facilitate a fine-grained and inclusive investigation of mobility-based segregation across a broad socioeconomic spectrum.

[1] Available at: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

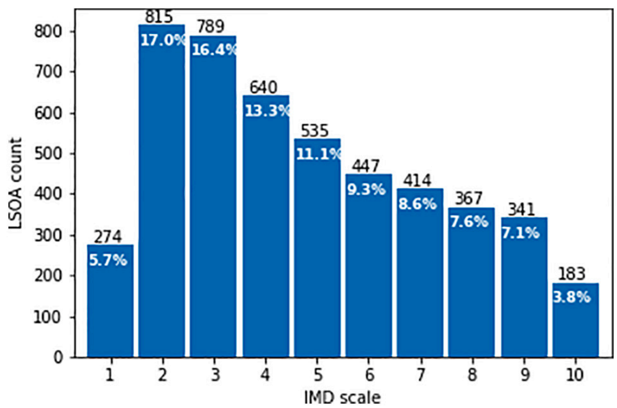

Fig. 1. Number of LSOAs per decile in London.

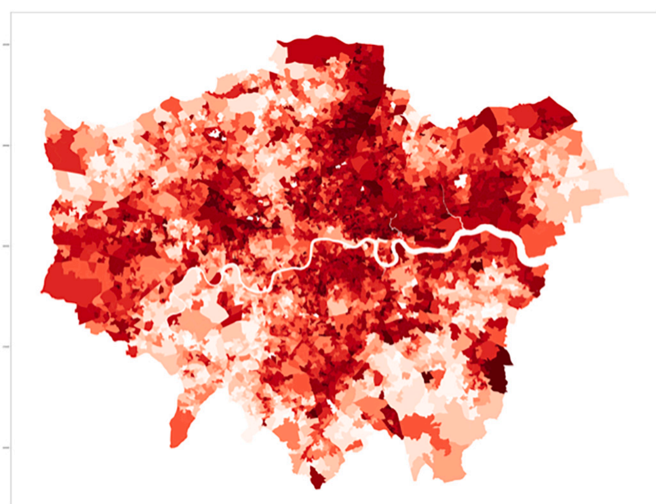

Fig. 2. Spatial distribution of IMD across London (more deprived LSOAs in darker colors).

4. Methodology

4.1. Measuring inter-LSOA mobility from Mobile network logs

Before mobility and segregation can be investigated, the mobile network log data must be processed into inter-LSOA trips. This entails joining the log data with LSOA data and then identifying trips that begin in a user's home LSOA and end in another LSOA. Then the frequency of trips between each pair of LSOAs can be tallied and analyzed.

4.1.1. Mapping between Antennas and LSOAs

The first step for measuring inter-LSOA mobility from mobile network logs is to identify the visits made between LSOAs. A challenge here is the mismatch between the boundaries of LSOAs and the coverage of the antennas. Depending on the deployment of the antennas, their coverage could be smaller than the boundary of one LSOA or it could span over multiple LSOAs. Therefore, a mapping is created between the antennas and the LSOAs by estimating the antennas' coverage and computing the overlap between the coverage and the LSOA boundaries. A Voronoi tesselation was used to estimate the coverage of antennas, which has previously produced reasonable results (Frias-Martinez, Williamson, & Frias-Martinez, 2011; Park, Serra, Martinez, & Nuria, 2018). Subsequently, it was observed that the Voronoi polygons frequently have a finer granularity than the LSOAs. For instance, Fig. 3 depicts the boundaries of LSOAs and the Voronoi polygons, and Fig. 4 shows the distribution of the size of the Voronoi polygons and LSOAs in square meters for the study area of London. The granularity of the Voronoi polygons suggests that it is feasible to locate mobile devices with respect to LSOAs based on the antennas that they connect to. In addition, both LSOAs and antennas are based on population density, such that small LSOAs are often associated with small antenna Voronoi polygons.

In some cases, an antenna coverage overlaps with a single LSOA, making it simple to link the two units. However, many antenna coverage polygons span multiple LSOAs (examples can be found in Fig. 3). In order to examine the impact of such cases on the study, it was first verified whether the LSOAs that overlap with an antenna coverage polygon are homogeneous in terms of their IMD decile. For each of the antennas coverage polygons, the standard deviation of the IMD decile for the overlapping LSOAs is computed. It is observed that the standard deviation is small in general; the distribution shows a highly positive skew, for example, 72% of the cases are below 1.5. We thus complete the

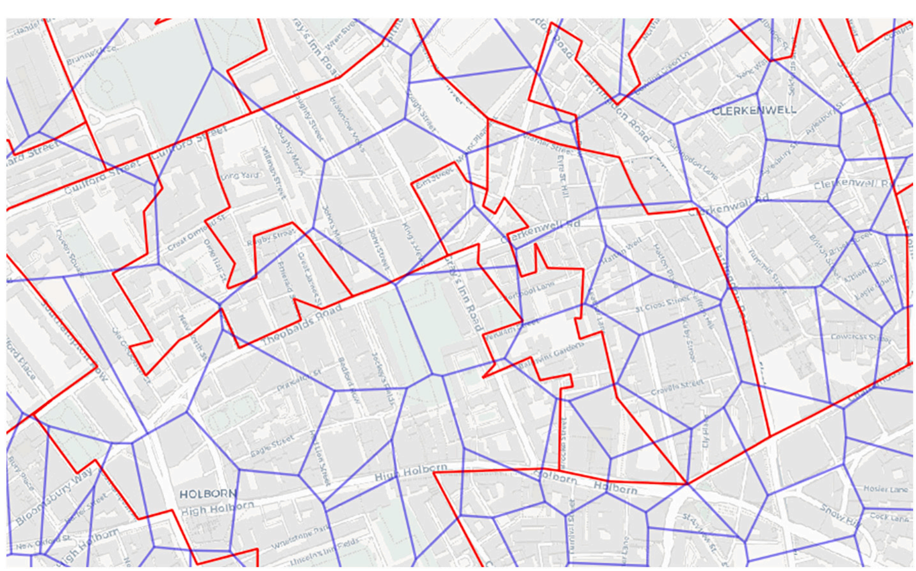

Fig. 3. Example boundaries of LSOA (red) and Voronoi polygons (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

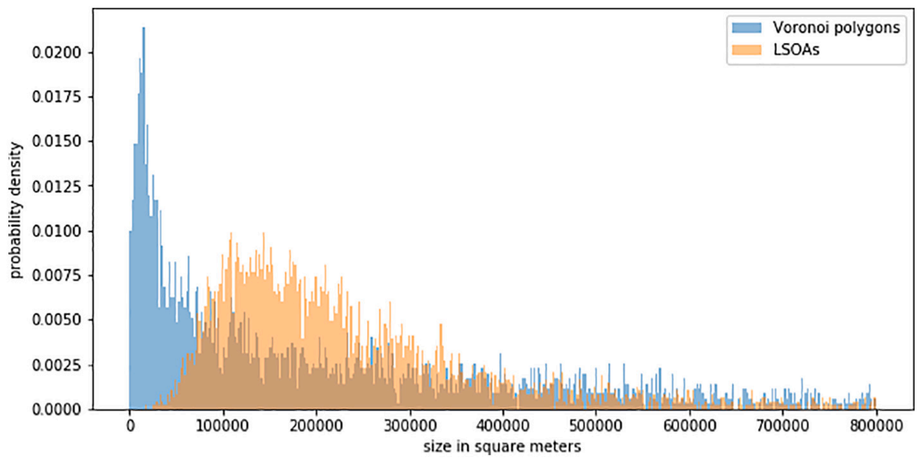

Fig. 4. Distribution of LSOA size based on area.

mapping simply by assuming that the mobility measured by an antenna is divided to LSOAs proportional to the size of the overlap between the polygon and the corresponding LSOA.

4.1.2. Home Detection

Methods for identifying residences from mobility traces commonly aim to identify routine activity during the evening and/or weekends (Bojic, Massaro, Belyi, Sobolevsky, & Ratti, 2015; Kevin, 2014). Likewise, a method for home antenna detection is developed in this context based on this intuition.

For each day over a month, devices that remain in the same antenna's polygon during the evening (midnight - 8 AM) for a substantial number of days are identified. The method first selects the set of antennas to which a device connects during the evening each day of the study. To ensure spatial accuracy, the radius of gyration (Appendix A) is employed to estimate the spatial deviation of a device and filter out the days when the estimated radius is larger than 2 km. Finally, the method classifies the home antenna of a device as the one which meets the above criteria for at least 14 days during the month, and the devices that do not have a home antenna that satisfies this criteria are excluded from the study.

Occasional hand-overs of stationary devices between distant antennas (also known as 'ping-pong handover') could decrease the accuracy of home detection; however, it is not likely that such handovers are prevalent in our dataset. Both the mobile device manufacturers and network operators have made efforts to reduce ping-pong hand-overs (e.g., (Neubacher, 2013; Tayyab, Gelabert, & Jäntti, 2019)), as it provides battery consumption benefits. In addition, in a dense deployment environment, such as London, the detection of home LSOAs are less sensitive to antenna handovers since the average size of the Voronoi polygons that approximate the coverage of antennas is 0.17 km2, which is smaller than the LSOAs average size of 0.32 km2. As a robustness check, the portion of person-day instances that are filtered by the 2 km spatial deviation threshold were measured. It was found that only 5% were filtered out during the midnight to 4 AM when minimal movement is assumed. Considering that these filtered instances may also include actual movements of the people in addition to ping-pong handovers, this implies that ping-pong handovers should not have a strong impact on home detection.

This method ultimately identifies the home antenna for 1.9 M devices, which are widely distributed across the city and overlap in coverage with 4813 LSOAs out of 4835. The results also had a large number of samples for each of the IMD deciles, for example, with a minimum number of 24,000 residents per decile.

4.1.3. Identification of Trips

For every device, all of the trips that are made to non-home LSOAs are extracted. This is achieved by identifying the connections to the antennas whose coverage does not overlap with the LSOA covered by the home antenna of the device. A 1-h time threshold for the duration of connection is also used in order to filter out potentially unstable connections that occur while mobile-phone users are on-the-move or in areas with weak connectivity.

4.2. Analyzing the relationship between urban mobility and segregation

An exploratory analysis of the extracted flows is first carried out using several visualizations to tease out the relationship between urban mobility and segregation. This includes the density of visits between IMD deciles and the distribution of distances between LSOAs by decile. These techniques provide initial evidence on the role of socioeconomic status and distance in generating urban mobility flows.

In order to further examine segregation-based mobility, it is theorized that there exists a socioeconomic distance-decay, whereby activity between LSOAs decreases as the IMD decile difference increases. This decay would be stronger if people have a tendency to restrict their daily mobility to locations with a more similar socioeconomic status.

A potential challenge to the socioeconomic distance-decay hypothesis is that mobility is simultaneously shaped by the geographical structure of LSOAs. The LSOAs of some classes might be physically closer to each other than those of the other classes (i.e., clustering due to place-based segregation), which could give rise to a traditionally theorized geographical distance-decay. In order to test the presence of a socio-economic distance-decay versus a geographical distance-decay, a gravity-type spatial interaction model is calibrated and their marginal effect sizes are compared.

Spatial interaction models are ideal for measuring the 'cost' or friction-of-distance factors between geographical regions, while accounting for the push and pull factors present in each region (Farmer & Oshan, 2017). A basic gravity-type spatial interaction model in its most general form is denoted by

$$T_{ij} = k\frac{V_i^{\mu} W_j^{\alpha}}{d_{ij}^{\beta}} \tag{1}$$

where T_{ij} is a matrix of flows between origins (subscripted by i) to destinations (subscripted by j), V represents origin attributes describing their emissiveness, W represents destination attributes describing their attractiveness, d is a matrix of the costs to overcome the physical separation between locations (usually distance or time), k is a scaling factor to be estimated to ensure the total observed and predicted flows are consistent, and μ, α, and β are parameters to be estimated that represent the effect of each variable on the flows. Since we are primarily interested in inference on the cost/distance variables, we can extend this basic model to a doubly constrained model that builds in information regarding the total outflow and inflow at each location in lieu of exogenous variables and is formulated as,

$$T_{ij} = A_i B_j O_i D_j d_{ij}^{\beta} \tag{2}$$

$$A_i = \frac{1}{\sum_j W_j B_j D_j d_{ij}^{\beta}} \tag{3}$$

$$B_j = \frac{1}{\sum_i V_i A_i O_i d_{ij}^{\beta}} \tag{4}$$

where O_i is the total number of flows emanating from origin i, D_j is the total number of flows terminating at destination j and A_i and B_j are balancing factors that ensures the total out-flows and total in-flows of the observed data are preserved in the predicted flows (Fotheringham & O'Kelly, 2020). This model can be calibrated using non-linear optimization where A_i and B_j are computed iteratively until convergence or by Poisson regression as suggested by (Flowerdew & Aitkin, 2020). The latter method is used here, which requires the logarithm of the cost/distance variables and a categorical variable for the origins and destinations (Tiefelsdorf & Boots, 1995). In the multiplicative model stated in (2–4), the distance term d is a multiplication of the two cost variables, $d_{geographic}$ and $d_{socioeconomic}$, each of which has its own parameter. Therefore, in the Poisson regression, which takes the log of the term d, the two are entered as separate additive independent variables[2] such that:

$$T_{ij} = exp\left(k + \mu_i + \alpha_j + \beta_1 lnd_{ij_{geographic}} + \beta_2 lnd_{ij_{socioeconomic}}\right) \tag{5}$$

where ln is the natural logarithm, and μ_i and α_j are origin-based and destination-based binary indicator variables that ensure the constraints enforced by the multiplicative balancing factors are met. The computation of the model was carried out using the *spint* module of the Python spatial analysis library (PySAL) and further details regarding model formulation, calibration, and interpretation are available in (Oshan, 2016).

5. Results

5.1. Exploratory visualizations

Using the individual mobility traces, a daily gyration measure was computed[3] for every user, which can be interpreted as the 'typical distance traveled'. It was then averaged across all days users are active over the period of the study. Fig. 5 reports the distribution of the measure as a boxenplot for each IMD decile where it can be seen that the median average gyration values tend to be greater for users that were identified as residents in LSOAs with higher IMD scores (i.e., less deprived). This result is consistent with some previously observed relationships between mobility and socioeconomic status (Dodson, Gleeson, & Sipe, 2004; Pappalardo et al., 2015). One reason for this result could be that individuals from different socioeconomic classes disproportionately use different modes of transportation. For example, people from more deprived areas are likely more dependent upon public transportation. Fig. 6 provides an aggregated overview of the public transport

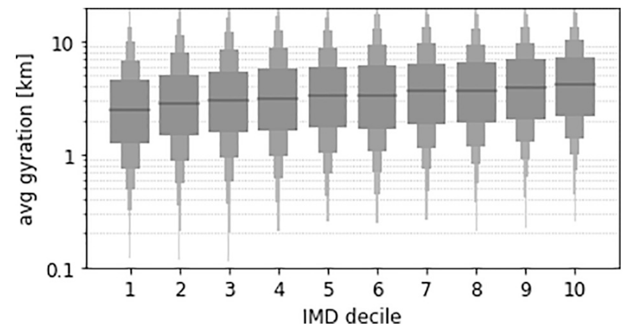

Fig. 5. Boxenplot of the average gyration.

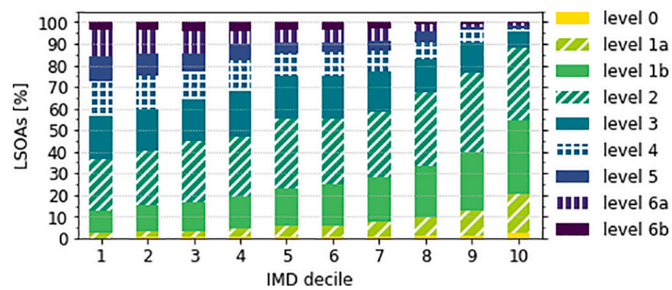

Fig. 6. Percentage of PTAL over IMD deciles.

accessibility level (PTAL) data[4] for London LSOAs, which measures the accessibility of a point to the public transport network based on walkability and service availability. Each LSOA is graded between 0 and 6b with a score of 0 is very poor access to public transport, and 6b is excellent access. In general, more deprived areas tend to have better access to public transportation while less deprived areas tend to have less access despite having larger average gyration values, suggesting that those living in less deprived areas have the resources to rely on private transportation. This discrepancy is not further pursued here as it is not the primary interest of this research, though the importance of this observation is highlighted later.

In order to obtain an overview of mobility flows between all IMD decile pairs, a visualization was constructed that consists of a 10 × 10 matrix where the rows and columns indicate the decile of the source and the destination, respectively (Fig. 7a). Furthermore, the matrix is normalized by rows (divide-by-sum) to account for the different number of LSOAs per decile. Consequently, the rows can be read as "going to" and each cell as the percentage of the visits made to the corresponding decile, with darker colors corresponding to higher percentages.

The rows of Fig. 7a show that the trips are dispersed throughout the IMD deciles, particularly across the middle of the decile distribution. This dispersion implies that the socioeconomic distance does not seem to act as a strong mobility deterrent (i.e., weak socioeconomic distance-decay). This pattern of dispersion is clearly contrasted when we juxtapose Fig. 7a with an IMD decile matrix created based on the neighboring frequency of LSOAs[5] (Fig. 7b). A stronger concentration along the diagonal in Fig. 7b shows that LSOAs of similar deprivation status frequently neighbor each other. Despite this segregated geographical layout of LSOAs, the dispersion of Fig. 7a indicates that people frequently visit beyond neighbor LSOAs that have larger socioeconomic distance. However, there still does exist a trend of fewer trips being made between long socioeconomic distances (e.g., from decile 1 to

[2] This implies a power distance-decay function, though we also tested an exponential distance decay function. The power distance-decay function was ultimately selected because it yielded better model fit in terms of root mean square error and pseudo r-squared

[3] the gyration is computed as a weighted measure based on the duration of antenna connections

[4] Available at: https://data.london.gov.uk/dataset/public-transport-accessibility-levels

[5] The neighboring LSOAs were computed based on the queen adjacency rule.

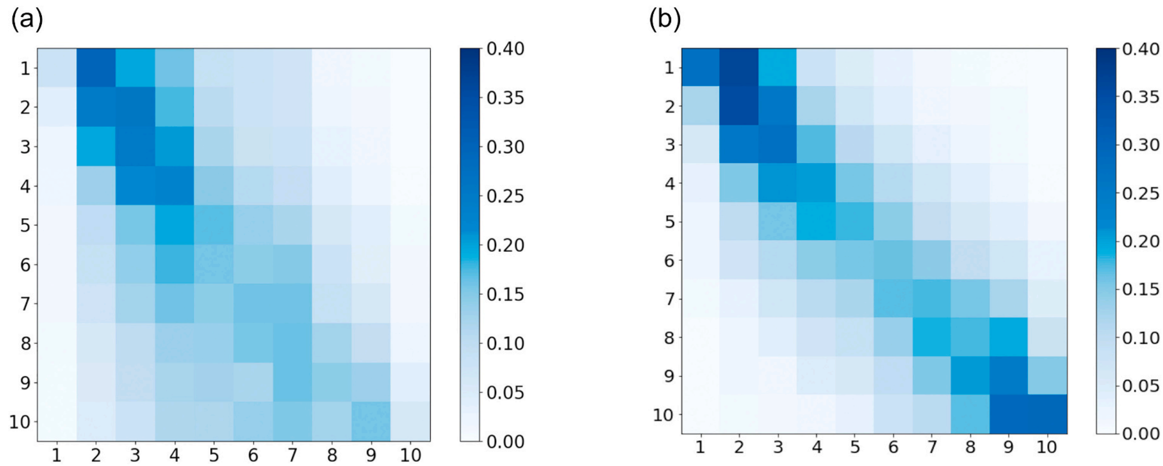

Fig. 7. Matrix Diagrams. Left shows the pairwise visits between IMD deciles and right shows the LSOA neighbors across IMD deciles.

decile 10 or vice versa).

A more holistic view is furnished through an analysis of the geographical distance distribution of subsets of LSOAs based on deciles. Fig. 8 contains a plot for each of the ten deciles; each plot visualizes the distributions of distances between LSOAs of one decile and each other decile, with lower IMD deciles distinguished by darker blue hues and higher IMD deciles distinguished by darker red hues. It is apparent that in general, distances tend to be lower amongst lower IMD deciles. This pattern becomes weaker as the plots move from IMD decile 1 to decile 5 and the distributions for IMD decile 5 to 7 are largely overlapping. The trend is reversed for IMD deciles 8 to 10 where there is a small but increasing proportion of short distances between higher IMD deciles. Overall, this suggests that LSOAs of high deprivation and low deprivation tend to be more nearby, but that high deprivation LSOAs (i.e., lower deciles) are the most clustered in space.

Overall, the visualizations presented here suggest a nuanced relationship between socioeconomic distance, geographical distance, and their impact on mobility flows. In order to further tease out these relationships, it is necessary to use multivariate techniques that can simultaneously account for different factors.

5.2. Spatial interaction models

The doubly constrained spatial interaction model was used to more formally investigate the impact of geographical distance and socioeconomic distance on mobility flows and quantify their individual distance-effects. More specifically, the number of trips made between each LSOA is employed as the dependent variable and the independent variables are composed of the cost/distance factors, including the difference between the IMD deciles of the source and destination LSOAs (i.e., socioeconomic distance) and the distance between the source and destination LSOAs centroids (i.e., geographical distance). Based on a Mcfadden's pseudo R^2 of 0.902, calibrating the described model provides a moderately-to-high fit to the data. The resulting parameter estimates are displayed in Table 1, along with standard errors and *t*-values for inference. However, it should be noted that the very large sample of more than 9 million flow

Table 1
Regression Coefficients.

Variable	Estimate	Std. Err	t-value
Intercept(k)	10.97	0.00060	18,108.54
log(Geographical Distance)	−1.97	0.00028	−3127.05
log(Decile Distance)	−0.03	0.0000030	−663,225.91

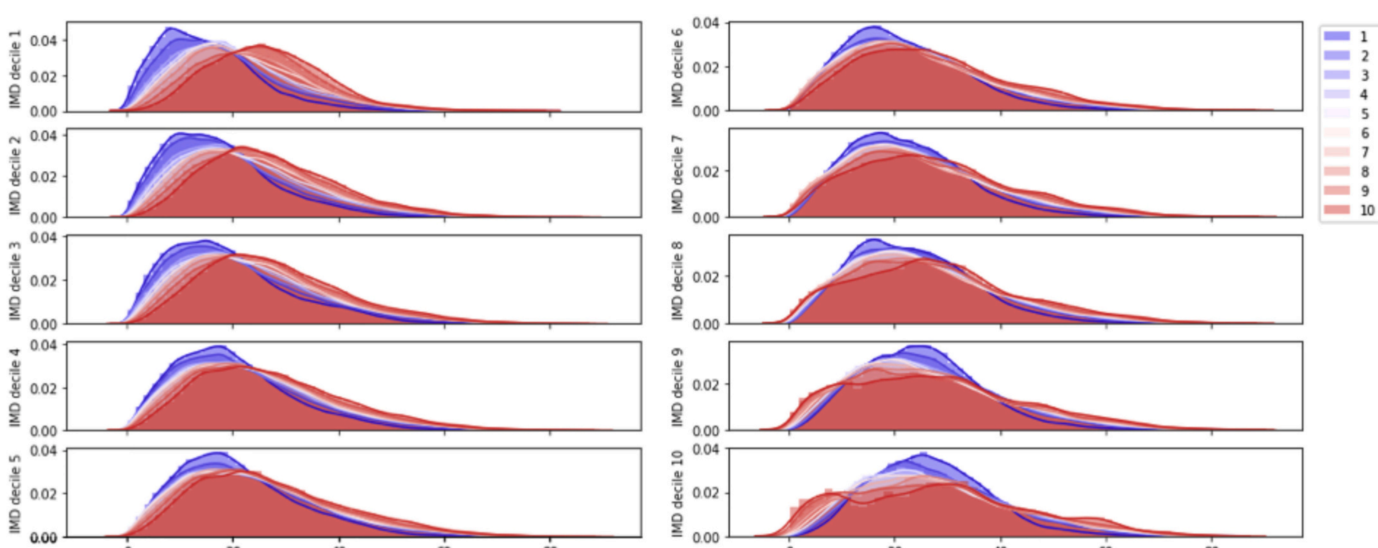

Fig. 8. Distribution of distance (in kilometers) to LSOAs of different deciles.

observation between the approximately 3000 LSOAs makes it relatively easy to obtain statistically significant results. Hence, the interest here is primarily on comparing the size and direction of estimated effect sizes, though small standard errors and large *t*-values indicate all estimates are significantly different from zero for standard levels of confidence.

The estimates for both geographical distance-decay and socioeconomic distance-decay take on the expected negative sign, indicating that the volume of trips decreases as either the geographical or socioeconomic distance increases. However, the magnitude of the estimate for socioeconomic distance-decay (−0.03) is relatively small, especially compared to the estimate for geographical distance-decay (−1.97). This means that although a socioeconomic distance-decay effect can be identified, it is minor compared to the association between geographical distance and the volume of trips. This suggests that people are more likely to choose to travel shorter distances to more diverse LSOAs than traveling over longer distances to more similar LSOAs. That is, they are more constrained in their daily activities by geographical distance than socioeconomic distance. These results are in agreement with the findings based on Figs. 7a that there is a weak socioeconomic distance-decay; however, this model is too simplistic. The decay effect could vary depending on the direction of a flow; for example, the effect could be stronger for flows from more deprived areas to less deprived areas than for those of the opposite direction. The results thus far also do not highlight differences observed in Fig. 8 across different portions of the socioeconomic spectrum. Mobility-based segregation is therefore further investigated by considering a possible interaction effect between socioeconomic and geographical distance and analyzing the flows by their directions.

In order to simplify the analysis, the 10 deciles are merged into 3 groups: the lower group (1–4), middle group (5–7), and the higher group (8–10). Then 9 directional categories are made based on inter-group movements. These categories are referred to as “L-L" (visits from a lower decile LSOA to another lower decile LSOA), “L-M", “L-H", and so on. Then the geographical distance as a continuous variable is interacted with these categorical variables. The outcome is that geographical distance-decay can be estimated individually for each of the directional categories. Without this three-way classification, there would be 100 directional categories, making the presentation and interpretation of results more complex. In addition, the classification is in alignment with patterns observed in Fig. 8 and was further corroborated using a machine learning approach that seeks to group LSOAs by simultaneously maximizing the similarity within groups and the differences between groups. In summary, a three-way logistic regression classification model was trained and tested that predicts the group of an LSOA based on its trip frequency to the 10 deciles. The model is implemented using scikit-learn (Pedregosa et al., 2011), and the class imbalance (different number of LSOAs amongst classes) is addressed by adjusting class weights (we set the ‘class_weight’ parameter to ‘balanced’). Table 2 presents the results of the classification model using a 5-fold cross validation, which achieved a reasonable level of classification performance (overall accuracy >75%).

The 9 categories based on the 3 classes were then used to create an interaction term with geographical distance that was entered into another doubly constrained interaction model and calibrated. The results are presented in Table 3 where as expected, geographical distance-decay is negative for all categories.

Table 2
Class Classification based on Mobility Destinations.

	Precision	Recall	F1	Num. of Items
Overall Avg. (Micro)	0.75	0.75	0.75	4813
Overall Avg. (Macro)	0.72	0.73	0.72	4813
Group “LOW”	0.83	0.85	0.84	2502
Group “MIDDLE”	0.64	0.57	0.60	1420
Group “HIGH”	0.69	0.76	0.72	891

Table 3
Regression Coefficients (Interaction between geographical distance and socioeconomic distance).

Variable	Estimates	Std. Err	t-value
Intercept	11.10	0.0006	18,423.0
log(dist)·H-H	−1.95	0.000012	−152,713.1
log(dist)·H-L	−2.06	0.000011	−188,076.2
log(dist)·H-M	−1.99	0.000013	−175,909.4
log(dist)·L-H	−2.07	0.000012	−173,115.6
log(dist)·L-L	−1.91	0.0000045	−424,264.1
log(dist)·L-M	−1.97	0.0000073	−271,358.0
log(dist)·M-H	−2.01	0.000012	−165,703.4
log(dist)·M-L	−1.98	0.0000072	−276,124.7
log(dist)·M-M	−1.94	0.0000078	−248,942.06

Moreover, several additional patterns emerge. First, it can be seen that distance-decay is the least negative (weakest) for the three intra-category groups (H-H, M-M, L-L), with the overall distance-decay being the weakest for trips from the low group LSOAs to other low group LSOAs. This means that on average trips are more likely to occur between physically distant LSOAs with similar socioeconomic standings than physically distant LSOAs with dissimilar socioeconomic standings and suggests that trips between more deprived LSOAs are the least deterred by geographical distance. This might indicate that residents in neighborhoods with lower socioeconomic status have a greater need to travel to another neighborhood to access amenities and employment opportunities than those in neighborhoods with higher socioeconomic status, while at the same time residents from both ends of the spectrum are more inclined to travel to more similar neighborhoods even if it entails longer distances. Second, distance-decay is moderate between groups that represent moderate socioeconomic distances (L-M, M-L, M-H, H-M) and the most negative (strongest) between groups that represent the largest socioeconomic distances (L-H and H-L). Together, these two trends support the hypothesis of a socioeconomic distance-decay in addition to a geographical distance-decay and that both types of distance-decays are more pronounced when they are combined.

6. Discussion

Several key findings from this research are highlighted below and some limitations and future research avenues are suggested.

6.1. Weak segregation in mobility: Bubbles can be broken

The first main finding from this research is that the socioeconomic indicator deployed in this study, the IMD, does not appear to have a strong segregating influence on mobility as trip flows were primarily constrained by geographical distance. This result suggests that individuals are more likely to make travel choices primarily regarding the distance of destinations than the socioeconomic status of destinations. In other words, people are not generally locked into socioeconomic bubbles in the context of mobility. A potential implication of this finding for urban planning and social inclusion policies is that segregation may be mitigated more efficiently by focusing on reducing the geographical distance-decay effect (i.e., cost to travel) in addition to various efforts to address socioeconomic inequality.

A reduction to geographical distance-decay could perhaps lead to an increase in societal integration. However, in order to assess this potential linkage it would be necessary to take advantage of a natural experiment using a scenario where distance-decay may become sensitive. For instance, a drop in fuel costs or the expansion of transport networks may decrease the constraints of geographical distance, stimulating mobility. In contrast, an increase in distance-decay could occur due to mobility restrictions. These changes in distance-decay may manifest disproportionately across the socioeconomic spectrum, which could be an interesting line of future inquiry. Importantly, to carry out such an investigation would require future research to incorporate more accurate measures of travel cost into the spatial interaction models in order to better capture potential sensitivities to the described interventions.

Another area to focus future research efforts would be to investigate whether the results obtained here vary within a study area and across study areas. As the results are based on a particular framework with a limited set of variables, a more detailed application of the result requires further investigation of diverse and complex socioeconomic and geographic factors. Variations within a study area could be investigated by decomposing a flow dataset by origin and individually analyzing trips from each origin to all other destinations (Fotheringham, 1981; Nakaya, 2001; Oshan, 2016). It would then be possible to reveal whether or not distance-decay varies spatially and if any potential variation is correlated with other factors, such as residential segregation, income, education, or heath disparities. Another approach would be to repeat the analysis for other cities and determine whether or not the evidence produced here is corroborated. However, it could be possible that urban areas with diverse social and spatial structures generate much different distributions of mobility flows. In addition, the calibration techniques employed here may become computationally prohibitive for a larger study area with many more spatial units, requiring alternative methods (Östh, Lyhagen, & Reggiani, 2016).

6.2. Bubbles are stronger for socioeconomically distant locations

The interaction effects observed between geographic distance and socioeconomic status provide a more nuanced understanding of mobility-based segregation. While visits decrease as geographical distance grows in general, the magnitude of the decrease is greater amongst the areas of greater socioeconomic distance (i.e, L-H and H-L flows), providing some evidence towards bubbles due to mobility-based segregation. Such a finding permits speculation on the the role of socioeconomic and geographical distance in the context of urban renewal or neighborhood revitalization projects, such as the establishment of a mixed-use complex or the enhancement of community infrastructure (e.g., (Scher, 2019)). Large projects often entail complicated trade-offs between location and the distribution and nature of benefits across many communities. Consequently, a useful strategy could be to consider investing in "bridging" projects that are located geographically to provide integration between communities that are socioeconomically distant. This may seem counter-intuitive in the case where conventional wisdom suggests directly investing in deprived communities. Such a counter-intuitive strategy may also be supported by recent work that approaches urban segregation from the perspective of spatial configuration; for example, Lageby's work (Legeby, 2013) exploring how structural urban patterns impact segregation in public spaces. Pursuing a combination of strategies would likely prove the most beneficial and would also contribute to restructuring the socio-spatial structure of society rather than solely treating the symptoms of systemic inequality. These findings may also be helpful for funding programs to develop and revise eligibility criteria. Many programs, such as e.g., scholarships (FSC, 2020), start-up grants (The Scout Association, 2020), and neighborhood renewal projects (OSCI, 2011)) use the IMD to deploy resources and it could be beneficial to target areas with relatively low mobility in addition to high deprivation.

An unexpected finding from this analysis is that there was near-symmetry for all pairs of opposite flow directions: distance-decay was similar between the flows for H-L and L-H, L-M and M-L, and also H-M and M-H categories. However, mobility flows from lower deciles to higher deciles (L-H) are inherently different from those from the higher deciles to lower deciles (H-L) due to different daily routines, access to services, and transportation modes. Though this analysis did not find evidence to support these hypothesized differences, the mobility-based segregation perspective is important because it allows directional asymmetries to be investigated and would not be possible through a traditional segregation lens based solely on static residential populations. The investigation of directional asymmetries in other contexts would therefore be an interesting avenue for future research.

6.3. Exploring the underlying mechanisms of mobility-based segregation

A promising direction for future research is to link the current findings with different urban disparities and further identify the underlying mechanisms of mobility-based segregation. One factor worth exploring is the distribution of the economic opportunities across the urban landscape and the formation of routine mobility patterns. A possible mechanism is that the location of employment and education opportunities that generate routine trips are not spatially aligned with certain socioeconomic groups, promoting socioeconomic mixing over shorter distances (i.e., weakening socioeconomic but not geographic distance-decay). The exploration of these economic factors would also provide insights regarding the spatial mismatch hypothesis (Kain, 2004) and whether those living in more deprived areas typically have longer commutes. Similarly, access to crucial services (e.g., health, groceries, transportation) is another factor worth examining by incorporating additional data sets about urban amenities (Daras, Green, Davies, Barr, & Alex, 2019; Office for National Statistics, 2020) and transport networks (Lucas, 2012). Furthermore, the exploration could be extended to interactions between such factors (e.g., between transportation network and economic opportunities) and enriched with information about local urbanism (e.g., East-West divide of London.) In this research, the identified trips do not include information regarding their purpose and future work could seek to use auxiliary data to incorporate contextual information to infer trip purpose (Sia-Nowicka et al., 2016) and link different mechanisms to different types of trips.

6.4. Advantages and disadvantages of the mobility data

Recall that in section 5.1 it was noticed that user average movement distance (i.e., gyration) was higher for less deprived areas despite being less accessible to public transportation. This observation suggests that attempts to analyze mobility solely based on public transportation could be limited because the data may systematically under-represent the mobility of some individuals. Therefore, a strength of the the data used here is that it is able to sample mobility activity across the entire population due to its inclusion of all antenna activity and high spatial coverage. This characteristic may have facilitated the finding of near-symmetry in the directional distance-decay estimates, which is in contrast to the results of Lathia et al. (Lathia et al., 2012) where asymmetries were detected based on socioeconomic status while analyzing public transportation logs. Another advantage of the our data that was not explored here is the high temporal resolution of the data. Recent work proposes using temporal subsets of data for time-specific spatial interaction models (Batty, 2018; Oshan, 2017; Oshan, 2020), which could be an interesting extension to the research presented here. For instance, this study could be replicated using daily, weekly, or monthly mobility flow matrices to investigate the potential evolution of geographical or socioeconomic distance-decay over time.

There are also some potential challenges to working with this type of data. First, the process of extracting trips between LSOAs based on antenna activity may be subject to some uncertainty. Future work could seek to validate trip flow volume against other sources of data or conduct sensitivity analysis for the various parameters used to define trips. In this case, the IMD was used as a proxy for socioeconomic status. However, future work could incorporate additional geodemographic variables such as race and ethnicity, which are commonly of interest in the segregation literature (e.g., (Wang et al., 2018)). Likewise, contextual information can also be incorporated from social media and land-use records. These endeavors could further enhance the potential to understand mobility-based segregation.

7. Concluding remarks

In this paper, segregation was approached through the lens of urban mobility and an inclusive analysis of the mobility between neighborhoods with different socioeconomic status was undertaken. Specifically, mobility-based segregation was investigated through the combination of radio access logs in London, UK, which capture extensive and diverse mobility activities across an urban population, with the IMD scores of small geographic areas to approximate socioeconomic status. Using these data sources overcame limitations from previous research that only examined relatively short time periods, a small number of socioeconomic classes, or a particular subset of the population. The presence of a socioeconomic distance-decay factor was then examined through the use of spatial interaction models, and found to have a subtle effect compared to geographic distance. Furthermore, it was observed that socioeconomic distance may interact with geographical distance, such that geographical distance-decay becomes stronger when the socioeconomic distance is greater.

The results obtained here demonstrate that by considering human mobility when inspecting segregation, a more nuanced picture emerges than what can be achieved based strictly on the residential perspective. An outcome is the ability to increase our understanding of how mobile populations can intensify or moderate the boundaries that define society, and to explore possibilities for promoting interactions across society. Some potential implications of these new insights were discussed in the context of planning and an array of extensions for this line of inquiry were outlined. This direction of work therefore warrants further attention and holds merit to contribute towards alleviating societal fragmentation and irreconcilable inequality.

Declaration of Competing Interest

None.

Appendix A. Appendix

Radius of gyration measures how far from the center of mass the masses are (Abramowicz, Miller, & Stuchlík, 1993). In the context of mobility, the measure has been used as an approximation of the average of traveled distances (Gonzalez, Hidalgo, & Barabasi, 2008). It is defined as the root mean squared distance between the set of antennas and its center of masses, weighted by the time spent with the connection. Formally:

$$g = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(t_j l_j - l_{cm}\right)^2} \tag{6}$$

where l_j represents the location of the j^{th} visited antenna, t_j represents the time spent in the j^{th} visited antenna and l_{cm} represents the location of the center of mass of the user's trajectory, calculated as $l_{cm} = \frac{1}{N}\sum_{j=1}^{N}\{t_j l_j\}$ and N the total number of antennas visited.

References

Abramowicz, M. A., Miller, J. C., & Stuchlík, Z. (1993). Concept of radius of gyration in general relativity. *Physical Review D, 47*(4), 1440.

Bach, P. B., Pham, H. H., Schrag, D., Tate, R. C., & Hargraves, J. L. (2004). Primary care physicians who treat blacks and whites. *The New England Journal of Medicine, 351*(6), 575–584.

Batty, M. (2018). Visualizing aggregate movement in cities. In *, 1753. Philosophical Transactions of the Royal Society B: Biological Sciences 373*. https://doi.org/10.1098/rstb.2017.0236 (Aug. 2018), 20170236.

Bojic, I., Massaro, E., Belyi, A., Sobolevsky, S., & Ratti, C. (2015). Choosing the right home location definition method for the given dataset. In *International conference on social informatics* (pp. 194–208). Springer.

Charles, C. Z. (2003). The dynamics of racial residential segregation. *Annual Review of Sociology, 29*(2003), 167–207. http://www.jstor.org/stable/30036965.

Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review, 106*(4), 855–902.

David M. Cutler and Edward L. Glaeser. 1997. Are Ghettos Good or Bad? *The Quarterly Journal of Economics* 112, 3 (1997), 827–872. doi:https://doi.org/10.1162/003355397555361 arXiv:/oup/backfile/content_public/journal/qje/112/3/https://doi.org/10.1162/003355397555361/2/112-3-827.pdf.

Dannamann, Teodoro, Sotomayor-Gómez, Boris, & Samaniego, Horacio (2018). The time geography of segregation during working hours. *CoRR abs, 5*(10). arXiv:1802.00117. http://arxiv.org/abs/1802.00117. https://royalsocietypublishing.org/toc/rsos/2018/5/10.

Daras, K., Green, M. A., Davies, A., Barr, B., & Alex, S. (2019). 2019. Open data on health-related neighbourhood features in Great Britain. *Scientific data, 6*(1), 1–10.

Dodson, J., Gleeson, B., & Sipe, N. G. (2004). *Transport disavantage and social status: a review of literature and methods*. Urban policy program. In *Griffith University Australia*.

Ellis, M., Wright, R., & Parks, V. (2004). Work together, live apart? Geographies of racial and ethnic segregation at home and at work. *Annals of the Association of American Geographers, 94*(3), 620–637.

Farmer, C., & Oshan, T. M. (2017). Spatial interaction. In *The Geographic Information Science & Technology Body of Knowledge (4th quarter 2017 edition ed.)*. John Wilson.

Farrell, C. R. (2008). Bifurcation, Fragmentation or Integration? The Racial and Geographical Structure of US Metropolitan Segregation, 1990—2000. *Urban Studies, 45*(3), 467–499.

Flowerdew, R., & Aitkin, M. (2020). *A Method of Fitting the Gravity Model Based on the Poisson Distribution. 22, 2 ([n. d.])* (pp. 191–202). https://doi.org/10.1111/%28ISSN%291467-9787/issues.

Fotheringham, A. S. (1981). Spatial structure and distance-decay parameters. *Annals of the Association of American Geographers, 71*(3), 425–436. http://www.jstor.org/stable/2562901.

Fotheringham, A. S., & O'Kelly, M. E. (2020). *Spatial interaction models:Formulations and applications*. Kluwer Academic Publishers. http://www.springer.com/earth+sciences+and+geography/geography/book/978-0-7923-0021-2.

Frias-Martinez, E., Williamson, G., & Frias-Martinez, V. (2011). An agent-based model of epidemic spread using human mobility and social network information. In *In Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on. IEEE* (pp. 57–64).

FSC. (2020). Opportunity Scholarship. https://www.field-studies-council.org/the-fsc-bursary-fund/.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *nature, 453*(7196), 779–782.

Jiang, Y., Li, Z., & Ye, X. (2019). Understanding demographic and socioeconomic biases of geotagged twitter users at the county level. *Cartography and Geographic Information Science, 46*(3), 228–242.

Kain, J. F. (2004). A pioneer's perspective on the spatial mismatch literature. *Urban Studies, 41*(1), 7–32.

Katz, B., & Lang, R. E. (2004). *Redefining urban and suburban America: Evidence from Census 2000*. Brookings Institution Press.

Kevin, S. (2014). Kung, Kael Greco, Stanislav Sobolevsky, and Carlo Ratti. 2014. Exploring universal patterns in human home-work commuting from mobile phone data. *PLoS One, 9*(6), Article e96180.

Krivo, L. J., & Peterson, R. D. (1996). Extremely disadvantaged Neighborhoods and urban crime. *Social Forces, 75*(2), 619–648. http://www.jstor.org/stable/2580416.

Kwan, M.-P. (2013). Beyond space (as we knew it): Toward temporally integrated geographies of segregation, health, and accessibility: Space–time integration in geography and GIScience. *Annals of the Association of American Geographers, 103*(5), 1078–1086.

Lathia, N., Quercia, D., & Crowcroft, J. (2012). *The hidden image of the City: Sensing community well-being from urban mobility Pervasive'12* (pp. 91–98). Berlin, Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-3-642-31205-2_6.

Lee, K. O. (2016). Temporal dynamics of racial segregation in the United States: An analysis of household residential mobility. *Journal of Urban Affairs, 2016*.

Legeby, A. (2013). *Patterns of co-presence: Spatial configuration and social segregation*. Ph. D. Dissertation. KTH Royal Institute of Technology.

Levinas, E. (1979). *Totality and infinity: An essay on exteriority* (Vol. 1) (Springer Science & Business Media).

Lindqvist, J., Cranshaw, J., Wiese, J., Hong, J., & Zimmerman, J. (2011). I'm the mayor of my house: Examining why people use foursquare-a social-driven location sharing

application. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 2409–2418). ACM.
Lucas, K. (2012). Transport and social exclusion: Where are we now? *Transport Policy, 20* (2012), 105–113.
Massey, D. S., & Denton, N. A. (1988). The Dimensions of Residential Segregation. *Social Forces, 67*(2), 281–315. http://www.jstor.org/stable/2579183.
Massey, D. S., Rothwell, J., & Domina, T. (2009). The changing bases of segregation in the United States. *The Annals of the American Academy of Political and Social Science, 626*(1), 74–90.
Mcgillion, M., Pine, J. M., Herbert, J. S., & Danielle, M. (2017). 2017. A randomised controlled trial to test the effect of promoting caregiver contingent talk on language development in infants from diverse socioeconomic status backgrounds. *Journal of Child Psychology and Psychiatry, 58*(10), 1122–1131.
Musterd, S., Marcińczak, S., van Ham, M., & Tammaru, T. (2017). Socioeconomic segregation in European capital cities. Increasing separation between poor and rich. *Urban Geography, 38*(7), 1062–1083. https://doi.org/10.1080/02723638.2016.1228371. arXiv.
Nakaya, T. (2001). Local spatial interaction modelling based on the geographically weighted regression approach. *GeoJournal, 53*(4), 347–358. https://doi.org/10.1023/A:1020149315435. April 2001.
Netto, V. M., Soares, M. P., & Roberto, P. (2015). 2015. Segregated networks in the city. *International Journal of Urban and Regional Research, 39*(6), 1084–1102.
Andreas Neubacher. 2013. Method for preventing ping-pong handover in mobile radio networks. US Patent 8,442,533.
Office for National Statistics. (2020). UK business; activity, size and location: 2019. https://www.ons.gov.uk/businessindustryandtrade/business/activitysizeandlocation/bulletins/ukbusinessactivitysizeandlocation/2019.
OSCI. (2011). Why the Indices of Deprivation are still important in the open data era. https://ocsi.uk/2011/03/24/why-the-imd-is-still-important-in-the-open-data-age/.
Oshan, T. M. (2016). *A primer for working with the spatial interaction modeling (SpInt) module in the python spatial analysis library (PySAL)*. *2* p. 11). https://doi.org/10.18335/region.v3i2.175.
Oshan, T. (2017). *A new era of spatial interaction: Potential and pitfalls*. PhD Dissertation. Arizona State University, Tempe, AZ.
Taylor M Oshan. 2020. Potential and Pitfalls of Big Transport Data for Spatial Interaction Models of Urban Mobility. *The Professional Geographer* 72, 4 (2020), 468–480. doi: https://doi.org/10.1080/00330124.2020.1787180 arXiv:doi:https://doi.org/10.1080/00330124.2020.1787180.
Östh, J., Lyhagen, J., & Reggiani, A. (2016). A new way of determining distance decay parameters in spatial interaction models with application to job accessibility analysis in Sweden. *European Journal of Transport and Infrastructure Research, 2*(2016), 16. https://doi.org/10.18757/ejtir.2016.16.2.3142.
John Östh, Ian Shuttleworth, and Thomas Niedomysl. 2018. Spatial and temporal patterns of economic segregation in Sweden's metropolitan areas: A mobility approach. Environment and Planning A: Economy and Space 50, 4 (2018), 809–825.
Pappalardo, L., Pedreschi, D., Smoreda, Z., & Giannotti, F. (2015). Using big data to study the link between human mobility and socio-economic development. *In 2015 IEEE International Conference on Big Data (Big Data).*, 871–878. https://doi.org/10.1109/BigData.2015.7363835.
Park, Y. M., & Kwan, M.-P. (2018). Beyond residential segregation: A spatiotemporal approach to examining multi-contextual segregation. In *, 71. Computers, Environment and Urban Systems*.
Park, S., Serra, J., Martinez, E. F., & Nuria, O. (2018). Mobinsight: A framework using semantic neighborhood features for localized interpretations of urban mobility. *ACM Transactions on Interactive Intelligent Systems (TiiS), 8*(3), 23.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*(2011), 2825–2830.
Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *Amer. Journal of Sociology, 116*(4), 1092–1153. http://www.jstor.org/stable/10.1086/657114.
Ioar Rivas, Prashant Kumar, and Alex Hagen-Zanker. 2017.Exposure to air pollutants during commuting in London: Are there inequalities among different socio-economic groups? Environment International 101 (2017), 143–157.
Sanbonmatsu, L., Ludwig, J., Katz, L. F., et al. (2011). *Moving to opportunity for fair housing demonstration program: Final impacts evaluation* (p. 2011). Washington, DC: US Department of Housing and Urban Development Office of Policy Development & Research.
Sandel, M. J. (2012). What money can't buy: The moral limits of markets. In *Macmillan*.
Scher, P. L. (2019). How Detroit became a model for urban renewal. *Fortune*. https://fortune.com/2019/01/02/detroit-cities-urban-renewal-innovation/.
Sia-Nowicka, K., Vandrol, J., Oshan, T., Long, J. A., Demar, U., & Fotheringham, A. S. (2016). Analysis of human mobility patterns from GPS trajectories and contextual information. *International Journal of Geographical Information Science, 30*(5), 881–906. https://doi.org/10.1080/13658816.2015.1100731. May 2016.
Silm, S., & Ahas, R. (2014). The temporal variation of ethnic segregation in a city: Evidence from a mobile phone use dataset. *Social Science Research, 47*(2014), 30–43. https://doi.org/10.1016/j.ssresearch.2014.03.011.
Smith, D. B., Feng, Z., Fennell, M. L., Zinn, J. S., & Mor, V. (2007). Separate and unequal: racial segregation and disparities in quality across U.S. nursing homes. *Health Aff Millwood, 26*(5), 1448–1458.
Smith-Clarke, C., Mashhadi, A., & Capra, L. (2014). Poverty on the cheap: Estimating poverty maps using aggregated mobile communication networks. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 511–520). ACM.
Tayyab, M., Gelabert, X., & Jäntti, R. (2019). A survey on handover management: From LTE to NR. *IEEE Access, 7*(2019), 118907–118930.
The Scout Association. (2020). HQ Start-Up Grants for new sections in deprived areas. https://members.scouts.org.uk/supportresources/4733.
Tiefelsdorf, M., & Boots, B. (1995). The specification of constrained interaction models using the. *SPSS loglinear procedure., 2*(1995), 21–38.
United Nations Centre for Human Settlements. (2001). *The State of the World's Cities, 2001. Vol. 27. UN-HABITAT*.
van Ham, M., & Tammaru, T. (2016). New perspectives on ethnic segregation over time and space. A domains approach. *Urban Geography, 37*(7), 953–962.
van Ham, M., Marcińczak, S., Tammaru, T., & Musterd, S. (2015). *A multi-factor approach to understanding socio-economic segregation in European capital cities*. Taylor & Francis.
van Kempen, R., & Murie, A. (2009). The new divided city: Changing patterns in European cities. In *, 4. Tijdschrift voor economische en sociale geografie 100* (pp. 377–398). https://doi.org/10.1111/j.1467-9663.2009.00548.x.
Wang, Q., Edward, N., Phillips, M. L., & Sampson, R. J. (2018). Urban mobility and neighborhood isolation in America's 50 largest cities. *Proceedings of the National Academy of Sciences, 115*(30), 7735–7740.
Wilson, W. J. (1990). *The truly disadvantaged: The Inner City, the underclass, and public policy*. University of Chicago Press. https://books.google.de/books?id=Z1Mz17ovNEYC.
Wilson, W. J. (2012). *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.
Zhang, X., & Wang, J. (2019). Mei-Po Kwan, and Yanwei chai. 2019. Reside nearby, behave apart? Activity-space-based segregation among residents of various types of housing in Beijing, China. *Cities, 88*, 166–180.