10-14-2016

Measuring urban segregation based on individuals' daily activity patterns: A multidimensional approach

Fei Li
Georgia State University, feili@gsu.edu

Donggen Wang
Hong Kong Baptist University, dgwang@hkbu.edu.hk

Recommended Citation
Li, Fei and Wang, Donggen, "Measuring urban segregation based on individuals' daily activity patterns: A multidimensional approach" (2016). *USI Publications*. 75.
doi: https://doi.org/10.1177/0308518X16673213

Measuring Urban Segregation based on Individuals' Daily Activity Patterns: A Multidimensional Approach

Journal:	*Environment and Planning A*
Manuscript ID	EPA-2014-0454.R1
Manuscript Type:	Original Manuscript
Keywords:	Urban segregation, activity space, a multidimensional approach, exposure, Hong Kong

Measuring Urban Segregation based on Individuals' Daily Activity Patterns: A Multidimensional Approach

Fei LI

Wagner Graduate School of Public Service, New York University, E-mail: fei.li@nyu.edu

Donggen Wang*

Department of Geography, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong. E-mail: dgwang@hkbu.edu.hk

Acknowledgement. The study is sponsored by two GRF grants from the Hong Kong Research Grant Council (HKBU245008 and HKBU247813).

Measuring Urban Segregation based on Individuals' Daily Activity Patterns: A Multidimensional Approach

Abstract. This paper develops a methodology to measure urban segregation based on individuals' socio-spatial experience of daily life. Since segregation can be considered as the isolation of people from those unlike themselves, its degree increases with the similarity in ethnicity, economic status, or other socio-demographic dimensions of interest between individuals and people who they are exposed to in their daily usage of urban space. Based on this perspective, we propose a regression estimator that measures segregation by assessing similarity or likeness between people and the social environments they experience in daily activity spaces. Compared to traditional segregation measures, the proposed estimator is not restricted to measuring residential segregation, but recognizes and assesses segregation as a dynamic process that unfolds in the daily life routines of individuals in a society and depends on the different ways individuals or social groups use urban space. It can be applied to various segregation factors, categorical or continuous, as well as to examine their interactions in a society. An empirical study in Hong Kong is used to demonstrate the proposed approach.

Keywords: urban segregation; exposure; activity space; a multidimensional approach; Hong Kong

1 Introduction

It has been increasingly recognized that segregation studies should go beyond residential place to daily activity space (Schnell and Yoav, 2001; Atkinson and Flint, 2004; Ellis, et al, 2004; Wang et al, 2012) and shift from location-based to people-based (e.g., Kwan, 2009; 2013). Krivo et al (2013) argue that conventional segregation studies, which consider residential

neighborhoods as the only context of social isolation, ignore the fact that social isolation likely extend from residential place to other places where people conduct out-of-home daily activity. Although residential place remains an important hub in individuals' daily life, the importance of other places (employment, recreation, etc.) has increased with the growth of human mobility in urban areas. Thus, a fuller understanding of urban segregation requires critical analyses of not only the socio-demographic compositions of residential neighborhoods, but also the types of social environments that individuals are exposed to in daily life.

Over the past few years a great deal of attention has been devoted to examine segregation or social isolation in individuals' daily activity space. For example, Schnell and Yoav (2001) use interactive and territorial spheres to examine individuals' exposure to other social/racial groups in daily life; Wong and Shaw (2011) use a location-based and activity-space-bounded approach to analyze individuals' exposure to different others; Wang Li and Chai (2012) study socio-spatial segregation through analyzing the similarities/ dissimilarities between individuals' activity spaces; Farber, Páez and Morency (2012) take a mobility-based perspective to study individuals' exposure to different linguistic groups in a bilingual society at both residential and non-residential neighborhoods.

These studies have successfully justified that segregation studies can be extended from residential neighborhood to daily activity space. Nevertheless, the understanding about social segregation/ isolation in daily activity space is far from complete. Among other issues, how to define and measure segregation based on individuals' socio-spatial experience in daily life is an imperative issue to be solved before the scholarship on activity space-based segregation can be established. As will be discussed in the next section, existing studies either use approaches that have unrealistic data requirements for broader application, or tend to overlook important individual differences in daily socio-spatial experiences. To contribute to this emerging literature, we propose a regression-based measure, which assesses segregation as the

correlation between the social character of individuals and that of the people to whom they are exposed in daily life circumstances. Since segregation is the sorting of people along certain social lines, the more the social environment an individual experiences in daily life matches her own identity, the more she is segregated. The proposed method is thus theoretically akin to the exposure/isolation index, yet less sensitive to the overall composition of the population as the latter does. It is also applicable to both categorically and continuously measured segregation factors. It allows the study of segregation in the entire daily life space or a part of it, as well as along different dimensions of segregation, including but not limited to income and race, and on their interactions. A case study in Hong Kong, using activity diary data collected from 770 individuals, is used to illustrate how the method works. We apply our method to three potential dimensions of segregation – age, income and private/public housing, and examine the interaction between different dimensions of segregation as well as the moderating effects of other factors.

The paper is structured as follows. The next section reviews the relevant literature. The third section introduces and discusses the proposed measure of segregation in daily activity space. The fourth section illustrates and validates the method using the Hong Kong case study. The last section concludes and discusses the strengths and limitations of the proposed approach.

2 From residential to activity space-based segregation: a literature review

Until the late 20th century, segregation studies had predominantly focused on the separation of racial or ethnic groups in residential neighborhoods, with the dissimilarity index proposed by Duncan and Duncan (1955) being the most popular measurement of residential segregation. The dissimilarity index is easy to calculate and has an intuitive interpretation – the proportion of members in a segregated group that needs to relocate to achieve a spatially even distribution.

Segregation has thus been implicitly defined as the uneven distribution of blacks and whites in most studies.

An alternative measurement that has gained increasing popularity in recent decades is the exposure/isolation index (Lieberson 1981), which originates from the idea that members of different groups can only interact with each other if they share the same neighborhoods. While each has its own advantages and disadvantages, the dissimilarity measure and the exposure measure represent two independent yet complementary aspects of segregation, defined respectively as the unevenness/dissimilarity dimension and the exposure/isolation dimension in Massey and Denton's (1988) five-dimensional framework of residential segregation[1], which have been later highlighted by researchers who argue they are the two most essential aspects of residential segregation out of Massey and Denton's five dimensions (Brown and Chung 2006; Reardon and O'Sullivan 2004).

The two segregation indices are not only technically, but also conceptually different, as Massey and Fischer (1999) observe:

> Whereas the dissimilarity index measures the extent of what might be called *structural* segregation between the two groups, the P* interaction index[2] captures more the *experience* of segregation from the viewpoint of the typical minority member. (p. 321-322)

While the former focuses on spatial distribution of social groups, the latter intends to capture the social distance and possibilities of interaction between social groups and is thus favored by some because it better reflects the original meaning of segregation (Johnston, Poulsen, and Forrest 2005). In this regard, exposure is not merely an alternative measurement of segregation; it implies an operational definition of segregation that substantially differs from the one that the dissimilarity index is based on.

---

[1] The "five dimensions" of segregation defined by Massey and Denton (1988) are actually five forms of different spatial distribution, which are conceptually different from the multiple dimensions of social segregation discussed later in this paper.

[2] The exposure index.

Studies that explore segregation or isolation in social network or activity space have mostly adopted this exposure-based perspective. For example, Echenique and Fryer (2007) develop a spectral segregation index based on the degree to which individuals are limited to same-race social connections. Lee and Kwan (2011) use a visualization approach to study individual's time-space potential to interact with others in their social networks. Farber, Páez, and Morency (2012) examine the isolation of English-speaking population, especially the elderly, from the French-speaking majority in Montreal, customizing each group's activity space with the estimated trip length of individuals. Farber, Neutens, Miller and Li (2013) exploit the land use and commuting patterns in a city to evaluate the potential that people can interact with each other after work. Krivo, Washington, Peterson and Kwan (2013) show that individuals who live in disadvantaged neighborhoods are also more exposed to high concentration of disadvantage in their daily activity destinations. While using different instruments and pursuing different research questions, these studies share an implicit assumption that exposure, or the potential to interact, in people's daily life matters – and to whom they are exposed to determines the level of segregation or isolation they experience.

Two studies of particular interest for the purpose of this paper among these are Schnell and Yoav (2001) and Wong and Shaw (2011), which attempt to construct a segregation measure in daily life spaces that is comparable to the conventional residential segregation indices. Schnell and Yoav (2001) construct the "everyday life space" along two axes:

> An interactive dimension designated by the social identity of meaningful others is measured along the three major spheres of daily activities outlined above: meetings with mates at work, meetings with friends, and telecommunication. A territorial dimension designated by four concentric territorial bases is measured spatially: close vicinity, cluster, neighborhood, and beyond. (p. 625)

They then calculate the exposure and isolation indices as the proportions of people that belong to the opposite and the same group of the individual within each sphere or territorial base, weighted by each sphere's importance (determined by both self-perceived importance and the amount of time spent in the sphere) in the individual's everyday life. Their measure is, therefore, essentially an individual-level version of Lieberson's (1981) exposure index, namely, each individual has a unique set of exposure and isolation indices.

A significant contribution of Schnell and Yoav (2001) is their definition of segregation as lying in the "ways of using space in the constitution of personal and social identities, rather than the residential agglomeration of social groups in neutral and static space" (p. 622). The introduction of the interactive dimension and the formation of socio-spatial isolation indices in non-geographical spheres also move beyond traditional approaches that are based on residential space and demographic patterns. Nevertheless, the territorial dimension is still residential-based, with the assumption that an individual has equal exposure to all locations within each home-centered concentric ring, which is hardly the case in everyday urban life. Moreover, evaluating isolation in the interactive spheres requires detailed personal interview about individuals' social networks, which inhibits wider application of the method than the illustrative example given by Schnell and Yoav (2001)[3].

Wong and Shaw (2011) refine the territorial bases with the concept of *activity space*, which refers to the space that individuals visit or utilize as they conduct everyday activities. While activity space is usually considered as an individual-level concept, Wong and Shaw aggregate it for each group in each residential neighborhood, hence creating a group-location based exposure index, which is arguably more applicable to large samples (though as Schnell and Yoav, they have to use simulated data to demonstrate it). Nevertheless, the aggregation

[3] The paper uses a hypothetical sample to illustrate how the measurement works. In a following case study on the segregation of African migrant workers in Tel Aviv, they interviewed sixty individuals and presented six representatives to reveal how the levels of segregation differ between migrant workers living in and outside a minority concentrated neighborhood.

procedure implicitly assumes that members of a certain group living in a spatial unit have similar activity spaces and therefore loses the very advantage of the activity space approach, i.e., acknowledging and preserving individuals' different daily socio-spatial experiences.

The exposure/isolation index and its variants, including the one developed by and Wong and Shaw (2011), measure segregation with the average proportion of "same"/"different" people, as defined by a binary variable (usually race), who reside in individuals' residential or activity space. While intuitively reflecting the aggregate level of segregation that individuals experience, the measure has two obvious limitations: it is restricted to measuring segregation by categorical variables, and, when the examined population group is small, the exposure to same group members will be inherently low, even if they are disproportionately isolated from those outside the group. The exposure index is hence seldom used alone and often complemented by the dissimilarity index, which, based on the different residential patterns of population groups, is not only restricted to categorical segregation factors but also hardly compatible with the idea of segregation in daily life space.

The conceptual difference between activity space segregation and residential segregation poses a number of challenges to the development of a proper measure. Firstly, instead of the single context of residential space, an activity space segregation measure needs to deal with multiple contexts in which daily activities such as work, shopping and socializing are performed (Krivo et al., 2013). Secondly, unlike residential patterns that are relatively stable, individual activity spaces and their sociodemographic characteristics (as defined by the people who are also occupying these spaces) are dynamic and hard to depict (Jones and Pebley, 2014). Thirdly, activity space segregation involves a time dimension because people visit activity destinations at different time points and spend varied lengths of time at these locations (Wang, Li, and Chai, 2012). Fourthly, activity spaces may have varied sizes and comprise different numbers of destinations, which inevitably affect individuals' exposure to other populations

(Wang, et al., 2012; Järv et al, 2014; Wang and Li, 2015). Fifthly, people may be present at the same place in different roles and/or for different purposes. For example, a waiter and a customer at a restaurant usually have very different social experiences even though they are exposed to the same group of people. Sixthly, mobility and travel is another important dimension that distinguishes activity space segregation from residential segregation. The use of different transport modes (e.g., private and public transport means) can lead to substantial differences in social exposure. Lastly, as segregation in activity space is essentially a personal experience, there is a question about how to aggregate individual exposure into citywide levels of segregation. A sensible measure of activity space segregation should address or at least acknowledge some of the challenges discussed here. This paper makes such an attempt and proposes an activity-space-based segregation measure that balances simplicity and the richness of individual activity patterns, which can be applied to both categorical and scale dimensions of segregation, as well as the interaction between two or more dimensions.

3 The measurement approach

3.1 An Individual-Based, Multidimensional Measurement of Segregation

Endorsing the exposure-based view of segregation, this study defines segregation as the isolation of people from those unlike themselves, or the tendency to gravitate to those like themselves, in their daily activities and use of urban space. Such “likeness” may be defined by various segregation factors, including but not necessarily limited to race/ethnicity and socioeconomics. Since it is the day to day encounter, interaction and assimilation people experience in daily life that establish and maintain their social positions and social distances with different others, the more the daily life environment of an individual matches her own social identities, the more she is segregated. Therefore, we measure activity space segregation

by the correlation between the social characteristics of individuals and those of their daily activity spaces.

The correlation measure, while based on individual usage of space, reflects the aggregate degree to which people gravitate towards those similar to themselves along various social dimensions, either categorical or continuous. The stronger the link between individual characteristics and the corresponding features of their environments, the higher the overall level of segregation in the city. To define the measure more formally, consider a single segregation factor, such as race. Let x_i denote the status of the ith individual on that factor, like whether or not the individual can be classified as a certain race, and y_i denote the concentration of similar people (in this case members of the same race) in the ith individual's activity space. The level of segregation by that factor can then be derived from a simple regression model:

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad (1)$$

The estimated coefficient β indicates the linkage between individuals' identities and those of the people that they encounter and potentially interact with in their daily life. For the overall level of racial segregation in a city or region, β would generally fall between 0 and 1. At one end of the spectrum ($\beta = 1$) is perfect segregation where everyone shares activity locations only with same-race people, while the other end ($\beta = 0$) represents perfect desegregation where an individual's race have little bearing on those of other people she meets on a daily basis. Statistically, if β is not significantly different from zero, segregation is minor or nonexistent. A positive and statistically significant β, on the other hand, would suggest that the studied factor is a significant dimension of segregation in the society.

The model may also be used to measure segregation within a certain population group (e.g., racial segregation among the elderly) and/or in a certain component of the activity space (e.g., the experience of racial segregation in shopping spaces). In these cases it is possible for β to be negative and significantly different from zero, which suggests that, instead of flocking towards

people of the same kind, the target group tend to visit urban locations characterized by people distinct from themselves (e.g., elderly people are more likely to visit destinations populated primarily by a different race than themselves), or that people tend to mingle with different others while performing certain activities (e.g., people prefer to shop in areas with more of other races). A negative β as an indicator of such "reverse segregation", however, should be interpreted only in relative terms. In our first hypothetical example, the elderly are less racially segregated than other age groups. In the second, shopping spaces are less racially segregated than other components of the activity space, such as work or recreational spaces. This is because, in order for β to be statistically significant, there has to be sufficient variation in y_i, or different concentration of races in activity spaces, which already suggests the existence of racial sorting or segregation[4]. In order for elderly minorities to visit white areas more frequently than minority areas, there have to be some "white areas" and "minority areas" to begin with, the distinction between which is presumably established by the larger numbers of younger people who also occupy these areas. Consequently, we do not expect a negative β when examining segregation for the entire population or in the activity space as a whole.

At this point, a discussion on the operationalization of activity space and y_i would be helpful. As in Wong and Shaw (2011), we define activity space as the subset of all urban locations that an individual visits as a result of his or her day-to-day activities (Golledge and Stimson 1997; Horton and Reynolds 1971). Information on these locations and activities can be derived from activity diary surveys, a data collecting tool that has been widely used in transportation and time use research (Kwan 2000; Stopher 1992; Wang, Chai and Li 2011). An

[4] This last statement might seem to suggest that the level of segregation in a society can be indicated by the variance of y alone, and the regression model is unnecessary. This is possible for residential segregation. However, in order to give an accurate account of activity space segregation, y has to reflect not only the social characteristics of those who reside in each neighborhood but also that of those who perform daily activities there, which would require information on every person's activity patterns in the city. The regression approach, in contrast, only requires information of a representative sample of individuals and their activity locations, and is more feasible under real-life data constraints.

activity diary survey typically asks respondents to recall and report all their activities and/or trips performed on one or several days, the destinations of which together constitute an individual's activity space. Social features of one's activity space can then be calculated as a weighted average of the corresponding characteristics of all locations in it:

$$y_i = \sum_j w_{ij} y_{ij} \quad (2)$$

where y_{ij} is the examined feature of the *j*th location in individual *i*'s activity space, and w_{ij} is the weight, or importance of the *j*th location in the *i*th individual's activity space, which may take account of the relative length of time the person spent there, the purpose and significance of the activity conducted there, and the person's role in that activity or place, with $\sum_j w_{ij} = 1$.

Depending on the studied segregation factor, y_{ij} can be measured in a number of ways. It may be derived from secondary sources, such as census data, or first-hand sources such as questions asked in activity diary surveys. Moreover, as mentioned before, it is well possible that people experience different levels of segregation in various domains of daily life, and model (1) and formula (2) may be applied to a subset of the activity space, e.g., destinations for social activities. In this sense, residential segregation is but a special case of activity space segregation, which is less dynamic and more visible than segregation in other parts of the activity space. When y_i is aggregated across all these parts, or, in other words, when the activity space is examined as a whole, the regression estimator β renders a more comprehensive account of the overall experience of segregation in people's daily life. With sufficient information, telecommunications, social networks, or the interactive sphere as defined by Schnell and Yoav (2001) may also be included as a component of the activity space.

The regression approach is also designed to study segregation along multiple dimensions, with standardized β used for comparing segregation levels between different dimensions to remove the effect of different scales and variances. By "multiple dimensions" here we are not referring to the various forms of differential spatial distribution discussed by Massey and

Denton (1988) or the various spheres of daily life spaces defined by Schnell and Yoav (2001); rather, we refer to the many social identities people hold that jointly define their positions in urban societies, or in other words, the multiple dimensions along which social divide and segregation may take place. Even in cities where one dimension of segregation dominates, it may still interact with other factors in significant ways. Two types of interactions are probably worth noting. One is the "reinforcing" effect between two or more dimensions of segregation, often correlated but not necessarily so, the joint force of which could multiply the experience of segregation for affected individuals. One example is race and income, as minority poor can be more vulnerable to both racial and income segregation (Jargowsky 1996). The other type of interaction is a "moderation" effect, where the moderator may not be a significant segregating factor by itself but affect the extent to which another factor segregates the society. For instance, car ownership or greater personal mobility may allow people to expand their activity spaces and hence have an offsetting effect on racial or income segregation, though people who do not have a car are not necessarily isolated from car owners.

To examine whether and to what extent two or more segregation factors reinforce each other, one may use a composite index of the segregation factors as the dependent variable or simply using one at a time. Take race and income again as an example. The model may take the following form:

$$y_i = \alpha + \beta_R x_{Ri} + \beta_I x_{Ii} + \beta_{R\times I} x_{Ri} x_{Ii} + \varepsilon_i \quad (3)$$

where y_i is a joint measure of the racial and income composition (or more broadly, a composite index of socioeconomic advantage/disadvantage) in the *i*th individual's activity space, x_{Ri} and x_{Ii} represent the individual's race and income, respectively, and $\beta_{R\times I}$ captures the additional effect of being both poor and a minority on activity space segregation.

The moderation effect can be similarly estimated by adding an interactive term to the model, though with no need to alter the dependent variable. For example, the following model tests whether car ownership (x_{Ci}) is an important moderator in racial segregation:

$$y_i = \alpha + \beta_R x_{Ri} + \beta_{R\times C} x_{Ri} x_{Ci} + \varepsilon_i \quad (4)$$

Assuming that the moderator (individual car ownership) is not directly correlated with the dependent variable (racial composition in activity spaces), the term $\beta_C x_{Ci}$ has been omitted from model (4). In empirical analysis it could be included to avoid possible confounding effects. The coefficient $\beta_{R\times C}$ will reveal in what direction and to what extent car ownership affects individuals' experience of racial segregation in daily lives.

3.2 The Regression Estimator and the Exposure/Isolation Index

It can be easily shown that, when applied to categorical segregation factors, the regression estimator β is both theoretically and structurally related to the isolation index, which is a reverse version of the exposure index and measures one group's exposure to its own members. The higher the isolation index, the more the group is segregated. Let *a* and *b* denote the two groups that divide a population of *N*, and x_i the group membership of the *i*th individual:

$$x_i = \begin{cases} 1 & \text{if the } i\text{th individual belongs to group } a; \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

A natural indicator of the concentration of group *a* at a location *j* would be the proportion of its members in all people who reside (or work/study/socialize/etc.) at that location, denoted by p_j. The isolation index of group *a* in activity space can then be defined as:

$$I_a = \frac{1}{n_a} \sum_i \sum_j w_{ij} p_{ij} x_i \quad (6)$$

where n_a is the number of individuals that belong to group *a*, w_{ij} is a weighing factor, and p_{ij} is the proportion of group *a* members in the *j*th location of the *i*th individual's activity space. x_i ensures that only the activity spaces of individuals who belong to group *a* will be

included in the calculation of I_a. Intuitively, this is the average proportion of same-group individuals in a group *a* member's activity space.

If we define y_{ij} in formula (4) as p_{ij}, and use the same weights w_{ij} in formulae (4) and (6), then

$$I_a = \frac{1}{n_a}\sum_i x_i y_i \tag{7}$$

Recall our simple regression model (1). In this case the regression estimator of inter-group segregation will be

$$\hat{\beta} = \frac{cov(x,y)}{var(x)} \tag{8}$$

or

$$\hat{\beta} = \frac{\sum_i x_i(y_i-\bar{y})}{\sum_i x_i(x_i-\bar{x})} \tag{9}$$

which is essentially I_a corrected for the mean of y_i and the variance of x_i. This correction is not merely a technical complication, but to adjust for the exposure/isolation index's underestimation of isolation for small population groups. Since the variance of x_i reaches its maximum when there is a 50-50 distribution between the two groups, and $\bar{y}$ controls for the average concentration of group *a* members in individuals' activity spaces, the regression estimator is more sensitive than the exposure/isolation index to segregation of small groups. Another advantage of the regression estimator over the exposure/isolation index, as mentioned before, is that it allows the estimation of segregation along multiple dimensions and their interactions, including both categorical and continuous factors.

3.3 The Properties of β

While our approach is fundamentally different from traditional residential segregation measures, especially the dissimilarity index, the regression estimator β possesses many of the properties that have been believed desirable for traditional segregation indicators. James and

Taeuber (1985) propose four criteria for segregation measures, originally discussed in the context of school segregation:

- *Organizational equivalence*: a segregation measure should remain unchanged if an organizational unit is divided into several units, each with the same composition of groups, or if several units with the same composition are combined into one. If, in the case that census tracts (or other spatial units) are used to characterize individuals' activity spaces, dividing a tract j into several structurally identical ones (i.e., with exactly the same characteristics or y_{ij}) will affect neither y_i nor the estimated $\hat{\beta}$.
- *Size invariance*: a segregation measure should not be affected if the number of individuals of each group in each unit is multiplied by a constant factor. Like organizational equivalence, this principle is also satisfied by the regression estimator.
- *The principle of transfers*: if an individual moves from a unit with a higher proportion of her group members to a unit with lower concentration of her group, segregation reduces, and vice versa. An analogous situation for activity space segregation would arise when an individual replaces one of the locations in her activity space that highly matches her own identity with one that matches her less. Intuitively, this will lead the corresponding y_{ij}, and subsequently y_i, to move in a direction that cause the regression estimator $\hat{\beta}$ to decrease, indicating a lower level of segregation.
- *Compositional invariance*: the segregation index should not be affected if the number of individuals of a certain group in each unit is multiplied by a constant factor. This refers to an overall change in the composition of the population. In James and Taeuber's definition, this is a desirable feature for "measures of dispersion" (e.g., the dissimilarity index), but not necessarily for "measures of central tendency" (e.g., the exposure index). As discussed before, the exposure/isolation index is actually strongly dependent on the overall composition of population, an issue that the regression estimator addresses by

controlling for the mean value of y_i and the variance of x_i. It would be difficult to formally prove compositional invariance of the regression estimator for two-group segregation, as the assumption of multiplying the members of a certain group would introduce extra individuals into the model, whose activity spaces may or may not be the same as existing members of that group. However, compositional invariance can be simply demonstrated for a continuous segregation factor x_i, such as income. If y_i is measured by the mean or median income in the *i*th individual's activity space, then a multiplication of everyone's income in the city will have same effects on x_i and y_i, leading to new income levels $y_i' = ky_i,\ x_i' = kx_i$. The modified model

$$y_i' = \alpha' + \beta' x_i' + \varepsilon_i' \tag{10}$$

would then have exactly the same $\hat{\beta}$ as model (1).

In sum, the regression estimator can be a rather simple and useful tool to study activity space segregation. Developed from the exposure definition of segregation, it provides more consistent measures of segregation levels than the exposure/isolation index when the population composition changes. It also allows the estimation of segregation along either categorical or continuous dimensions, as well as the reinforcing effects between different dimensions of segregation and the moderating effects of other factors. Moreover, the data requirement for this method is quite flexible. While it is completely applicable with readily available information such as population census, better data sources or data collection techniques that might emerge in the future could greatly enhance the potential of this method in segregation research.

4 An empirical application of the measure

4.1 Case and Data

A web-based activity diary survey was conducted in Hong Kong, China between July and November, 2010, providing information on 770 individual's activity spaces on a typical weekday. The activity space comprises the destinations of all activities but those performed at home, as people are not exposed to potential social interactions, except for arranged meetings and interactions among immediate family members, when they are in their homes. The neighborhoods where they live, nevertheless, can be and usually are part of their activity spaces, if they performed some out-of-home activities in their home neighborhoods.

Table 1 shows the sample profile and compares it to the general population in Hong Kong. Due to the use of an online survey, the sample is considerably better-educated and has a younger age distribution than the overall population. It also contains larger shares of female and single individuals, and a smaller share of people from low-income households. As for housing type and the region of residence, nevertheless, the distribution of respondents closely matches that of the general population.

Despite the city's colonial history, ethnicity is hardly a dominant factor in Hong Kong as in traditional Western segregation research, as the population is predominantly Chinese, with a small group of transnational elites and a community of foreign (mainly Philippine and other Southeast Asian) domestic workers. Social divides along economic factors are more visible, including that established by one of the largest social housing system in the world, which accommodates roughly half of the population in Hong Kong (Lo, 2005). The few studies on socio-spatial segregation in Hong Kong argue that public housing in Hong Kong has had an important integrative role (Delang and Lung 2010; Forrest, La Grange, and Yip 2004), though the dissimilarity indices of public housing versus private housing residents are found to be high (Forrest, La Grange, and Yip 2004), and the usually large public housing estates themselves may constitute enclaves of poverty at a micro level.

Considering the contexts, we test our measurement approach on three dimensions of segregation - age, income, and the private/public housing divide in Hong Kong. This is not meant to be an exhaustive list of socio-spatial divides in Hong Kong; nor are we claiming that these are the three most important ones. The aim here is more to demonstrate how the proposed measurement approach works and how different dimensions of segregation co-exist and interact, rather than to develop a comprehensive narrative of the socio-spatial structure of Hong Kong. To show how interactions between segregation factors can be examined, we further examine the potential reinforcing effect between income and housing segregation and the moderation effects of personal mobility on the three types of segregation.

The three key individual variables (x_i) are age, monthly household income, and housing type. While housing type is a binary variable (living in private housing = 1), the other two have been collected using Hong Kong Census categories and treated as continuous variables. Age contains seven categories, from 15 to 75. Household income has eleven categories, from 5,000 to 105,000 Hong Kong Dollars (or roughly 645 to 13,500 US dollars) per month. The corresponding destination variables, y_{ij}, are the median age, median household income, and percentage of households living in private housing in each neighborhood[5] derived from the 2011 Hong Kong Census. In calculating activity space characteristics (y_i), they are weighted by the time an individual spent at each location during the diary day.

<Insert Table 1 about here>

4.2 The Three Dimensions of Segregation

[5] The spatial unit used in this study is Grouped Constituency Area (GCA), which has been used in both the activity survey (to record where individuals live and perform activities) and the calculation of population statistics. It is based on the District Council/Constituency Area system in Hong Kong, which has been established for district administration and election affairs. To facilitate the reporting of activity diaries, we grouped the 390 Constituency Areas into 109 GCAs according to commonly perceived boundaries of broader neighborhoods.

Table 2 shows the levels of age, income, and private/public housing segregation in Hong Kong estimated by the regression method. To better illustrate individuals' experience of segregation in the various domains of daily life, we contrast segregation levels in their residential and employment spaces, for which the dependent variable y_i represents the characteristic of individuals' residential or working places, to that in the activity space as a whole. Table 2 presents standardized βs to remove the effect of different scales when comparing segregation levels between different social dimensions.

From what Table 2 shows, age is clearly a weak segregation factor in Hong Kong, while people are systematically sorted by income and housing type in their daily lives. Income segregation is much stronger in residential space than in other parts of the activity space: one standard deviation change in an individual's household income is, on average, associated with a 1/3 standard deviation change in the median household income of her residential neighborhood, but only a 7.7% standard deviation change of the median household income in her entire activity space. Housing appears to be the most important dimension of segregation among the three, and the type of housing an individual lives in closely predicts the composition of private/public housing residents not only in her residential space but also in the rest of her activity space. The proportion of private housing residents in the activity space of an average private housing resident is 16.5% higher than that of someone who lives in public housing.

<Insert Table 2 about here>

4.3 The Reinforcing and Moderation Effects

While income and the type of housing one lives in are clearly correlated, they are not necessarily equivalent in their implications for segregation. Since there are always more applicants than available units, not all low income families can be accommodated in public

housing. More importantly, living in private or public housing may also foster different lifestyles and social networks, as well as different accessibility to facilities and community services. It would be of interest, therefore, to test the interaction between the two segregation factors. If a "reinforcing" effect as discussed above is at play, poor families living in public housing may be particularly isolated for suffering from both income and housing segregation.

To that end, we employ model (2) with y_i as a joint index, calculated by averaging the z scores of the median household income and the proportion of private housing residents in a neighborhood. It is supposed that the two variables work in the same direction, as private housing residents generally have higher income than public housing residents. Table 3 presents the estimation results. Somewhat unexpectedly, when housing type is controlled for, income is no longer significant in determining the social environment one performs daily activities in. Neither is the interaction term between income and housing type. In this case, income may not represent an independent factor of segregation, but merely reflects part of the deep-rooted private/public housing divide in Hong Kong. If the coefficients of the interaction terms are significantly positive, though, a "dual segregation" might be present, further marginalizing the disadvantage group who are both low-income and living in public housing. If, on the contrary, the interaction coefficients are negative and significant, then the public housing system might have played an integrative role and mitigated income segregation for its residents.

<Insert Table 3 about here>

The other type of possible interaction – the moderation effect – is illustrated using personal mobility as an example. Personal mobility is assessed by private car ownership (*car*=1 when the individual's family owns 1 or more cars, 0 otherwise) and the size of individual activity space (measured by the standard distance between all activity destinations, *sd*), as people who

have limited mobility or small activity spaces can be bounded to more segregated environments. Since preceding analyses have shown that private/public housing might be the only independently significant segregation factor among the three, we only examine the effect of mobility on housing segregation. Table 4 shows the results. Both car ownership and the size of activity space are found to have negative effects on housing segregation, suggesting that individuals with higher mobility are less likely to be isolated by housing status. The effect of car ownership is much stronger than that of the size of activity space, the magnitude of which is almost as large as the segregation levels in workplace or activity space as a whole, suggesting that owning a private car could offset a great part of the isolating effect of living in public housing. However, private car ownership rate in Hong Kong is much lower than in similar megacities across the world. By 2010, the number of registered private cars per household in Hong Kong is below 0.19. Only 18.8% of the respondents in our survey are private car owners, and 73.8% of them live in private housing. The ownership of private cars often indicates socioeconomic privilege, which also explains why *car* is significantly correlated with the composition of private/public housing residents in activity spaces.

The size of activity space (*sd*) is less directly related to socioeconomic status, though the variable has a significant positive effect on the proportion of private housing residents in activity space, suggesting that people who have wider horizons in daily usage of urban space are more exposed to private housing residents. Larger activity space also has an offsetting effect on private/public housing segregation, though the effect is relatively small compared to the strong tendency towards housing segregation in Hong Kong, and mostly insignificant apart from in residential space. Nevertheless, the results imply that personal mobility could be an important moderator in urban segregation, the effect of which may be more accurately captured by using an instrument that better controls for socioeconomic status. In a city highly depending

on public transit like Hong Kong, the placement of public housing and access to public transit could greatly affect the daily experience of segregation by public housing residents.

<Insert Table 4 about here>

4.4 Subgroup Tests

It is possible to estimate and compare the daily experience of segregation by different groups in a city. Since the Hong Kong sample over represents young, well-educated individuals and females, subgroup analysis by age, education attainment and gender may also serve as tests for potential sampling bias in estimated overall segregation levels. For illustration purposes, we focus on private/public housing segregation as in the previous analysis. Table 5 shows the results of subgroup tests.

The top panel of Table 5 compares housing segregation between three age groups. There is not much difference in terms of residential segregation; if anything, the elder group appears to be slightly less segregated in residential space than younger people. In the whole activity space, however, the elder group is almost twice as likely to be exposed to people living in the same type of housing that they themselves live in as the younger groups. This is possibly because elder people have smaller and more home-bound activity spaces. The middle-aged group is the only group that is significantly segregated by housing type in work space, which could be due to the higher percentage of individuals that have regular workplaces in this group. The underrepresentation of the elderly in our sample, therefore, might have led to an underestimation of housing segregation in the activity space.

<Insert Table 5 about here>

Likewise, the subgroup test by gender shows that men are slightly more segregated by housing type in residential space, whereas women experience significantly higher housing segregation in workplace and the activity space as a whole. As for individuals with different education levels, no substantial differences are detected in residential or activity space segregation, though people with college or higher degrees are more sorted by housing type in workplace than those with high school or lower degrees. If women and elder people are more vulnerable to activity space segregation owing to limited mobility or disadvantage, workplace segregation of the higher education group probably more reflects self isolation, in the sense that workers with higher human capital have greater freedom to choose jobs in environments that cater to their own social identities.

4.5 Summary

Using the activity diary survey of 770 individuals in Hong Kong, we have illustrated the application of the regression method to daily life segregation along various social dimensions. We find varied levels of socioeconomic segregation in individuals' activity spaces, which would be overlooked by the conventional approach that centers on residential segregation. Among the three segregation factors we have examined, the private/public housing divide appears to be the strongest in all three contexts. Income is a significant segregation factor when tested on its own, though income segregation in Hong Kong usually coincides with housing segregation and could be largely explained by the latter. We find age a weak segregation factor compared to the other two, but elder people are more prone to socioeconomic segregation in their activity spaces.

Personal mobility, including private car ownership and travel distance or the size of activity space, is found to significantly mitigate the level of segregation people experience. The moderation effect applies not only to activity space segregation, but to residential segregation

as well, implying that higher mobility might allow individuals to choose more freely where they live (or, conversely, individuals that choose not to live in enclaves of their own class may have to travel longer or to purchase private cars). The economic privilege associated with private car ownership in Hong Kong could have accentuated this effect, which may or may not be found in cities with high car ownership levels, though we expect that segregation in activity space is closely related to individuals' capacity to travel to and make use of different urban locations.

5 Discussion

A common feature of large cities nowadays is the great diversity – not merely the diversity in ethnicity, but the different cultures, identities, social classes, and lifestyles that the cities embrace. Diversity is a key element of what makes cities dynamic, vibrant, and attractive; yet it also creates various divisions along which the urban society can be fragmented and segregated. This paper makes an attempt to develop a new approach that may help extend the conventional understanding of segregation to better address the growing diversity and the various forms of segregation in contemporary cities. The approach is based on individuals' experiences of isolation in daily life and measures segregation by the extent to which people live, work, and perform daily activities in urban environments that match their own social identities.

The proposed measure is theoretically akin to the exposure/isolation index, yet more suitable for comparison across different cities or groups than the latter, which is highly dependent on the overall population composition. The new measure is also flexible enough to be applicable to both categorical and continuous dimensions of segregation. With a representative sample and a well-designed survey, the method can produce simple measures of citywide segregation in the whole activity space or in various domains of daily life, as well as

along multiple social divides. Moreover, it allows easy analysis of the interaction between different segregation factors and the moderation effects of other factors, which could help researchers and policymakers determine the driving forces of segregation, locate the most marginalized group, and devise more effective anti-segregation measures.

One caveat of the regression approach is that, when the segregation factor studied is a continuous variable, such as age or income, using the median or mean value of the factor as the dependent variable ignores its variation within neighborhoods or activity spaces. In other words, we assume that the median or mean value represents the overall character of the activity space on that factor. This simplification may affect the accuracy of the measure, especially if 1) the spatial unit of measurement is too large so that the social feature of a neighborhood can hardly be defined by measures of central tendency, 2) some individuals' activity spaces include places with distinct characters, or 3) the variance of the studied variable between neighborhoods is relatively small compared to that within neighborhoods. In the last situation, the studied variable is likely not a meaningful segregation factor, like age in the Hong Kong case, thus the analysis is less interesting in itself. The first problem can be addressed by using finer spatial units, or other ways to measure social environment in place of or complement to census data, such as self-reported socio-spatial experiences collected in the activity diary (for example, see how Vittengl and Holt 1998 measure mood in diary reports or the geo-narrative approach proposed in Kwan and Ding 2008). This could also help circumvent the limitations of traditional segregation indices due to the arbitrary sizes and boundaries of census tracts (Cortese, Falk, and Cohen 1976) and better target the people that an individual actually shares daily activity space and interacts with, thus producing more accurate segregation measures.

Situation (2) is worth some extra discussion. When an individual routinely visits very different places (e.g., living in a poor neighborhood while working in a rich one), using the median or mean value of the studied variable to describe her activity space may obscure the

actual composition of people she is exposed to everyday. An easy workaround is to categorize the continuous variable and then separately test the level of segregation for each stratum. This can be helpful if one is interested in the isolation of certain groups, such as the very rich or very poor. For a general estimation of income segregation in a city, nevertheless, we do not recommend this approach, as arbitrary categorization can be more problematic than oversimplifying the character of activity space by measures of central tendency. If the purpose is to build an individual-level segregation measure that assesses the degree of isolation a person faces in her daily life, the composition of people in her activity space should definitely be a key factor in it. In order to assess segregation at an aggregate level, more summary measures of individual experience are usually sufficient, for it is the collective pattern of individual choices and experiences that we seek to measure. Having a diversified activity space *per se* indicates a personal choice against segregation. If the majority of people have similar diversity in their activity spaces, the overall level of segregation should be low, as will be truly reflected by a small or insignificant β due to the lack of variance in y_i or correlation between x_i and y_i.

References

Atkinson R, Flint J, 2004, "Fortress UK? Gated communities, the spatial revolt of the elites and time–space trajectories of segregation" *Housing Studies* **19** 875-892

Blumen O, Zamir I, 2001, "Two social environments in a working day: occupation and spatial segregation in metropolitan Tel Aviv" *Environment and Planning A* **33** 1765-1784

Brown L A, Chung S Y, 2006, "Spatial segregation, segregation indices and the geographical perspective" *Population, space and place* **12** 125-143

Cortese C F, Falk R F, Cohen J K, 1976, "Further considerations on the methodological analysis of segregation indices" *American Sociological Review* **41** 630-637

Delang C O, Lung H C, 2010, "Public housing and poverty concentration in urban neighbourhoods: The case of Hong Kong in the 1990s" *Urban Studies* **47** 1391-1413

Duncan O D, Duncan B, 1955, "A methodological analysis of segregation indexes" *American sociological review* **20** 210-217

Echenique F, Fryer R G, 2007, "A measure of segregation based on social interactions" *The Quarterly Journal of Economics* **122** 441-485

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** 620-637

Farber S, Páez A, Morency C, 2012, "Activity spaces and the measurement of clustering and exposure: a case study of linguistic groups in Montreal" *Environment and Planning-Part A* **44** 315

Farber S, Neutens T, Miller H J, Li X, 2013, “The social potential of metropolitan regions: A time-geographic measurement approach using joint accessibility” *Annals of the Association of American Geographers* **103** 483-504

Forrest R, La Grange A, Yip N M, 2004, "Hong Kong as a global city? Social distance and

spatial differentiation" *Urban Studies* **41** 207-227

Golledge R G, Stimson R J, 1997 *Spatial behavior: A geographic perspective* (Guilford Press , New York)

Horton F E, Reynolds D R, 1971, "Effects of urban spatial structure on individual behavior" *Economic Geography* **47** 36-48

James D R, Taeuber K E, 1985, "Measures of segregation" *Sociological Methodology* **14** 1-32

Jargowsky P A, 1996, "Take the money and run: Economic segregation in US metropolitan areas" *American sociological review* **61** 984-998

Järv O, Ahas R, Witlox F, 2014, "Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records" *Transportation Research Part C: Emerging Technologies* **38** 122-135

Jones M, Pebley A R, 2014, "Redefining neighborhoods using common destinations: Social characteristics of activity spaces and home census tracts compared" *Demography*, **51**(3) 727-752

Johnston R, Poulsen M, Forrest J, 2005, "On the measurement and meaning of residential segregation: a response to Simpson" *Urban Studies* **42** 1221-1227

Krivo L J, Washington H M, Peterson R D, Kwan M P, 2013, "Social isolation of disadvantage and advantage: The reproduction of inequality in urban space" *Social Forces* **92** 141-164

Kwan M P, 2000, "Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: A methodological exploration with a large data set" *Transportation Research C* **8** 185-203

Kwan M P, Ding G, 2008, "Geo-Narrative: Extending Geographic Information Systems for Narrative Analysis in Qualitative and Mixed-Method Research" *The Professional Geographer* **60** 443-465

Lee J Y, Kwan M P, 2011, "Visualisation Of Socio-Spatial Isolation Based On Human Activity Patterns And Social Networks In Space-Time." *Tijdschrift voor economische en sociale geografie* **102** 468-485

Li F, Wang D, 2014, "Cyberspace: Connected or Segregated? From Digital Divide to Virtual Segregation" *Environment and Planning B* **41** 323-340

Lieberson S, 1981, "An asymmetrical approach to segregation", in *Ethnic segregation in cities* Eds C Peach, V Robinson, S Smith (Croom Helm, London)

Lo C P, 2005, "Decentralization and polarization: Contradictory trends in Hong Kong's postcolonial social landscape" *Urban Geography* **26** 36-60

Massey D S, Denton N A, 1988, "The dimensions of residential segregation" *Social forces* **67** 281-315

Massey D S, Fischer M J, 1999, "Does rising income bring integration? New results for blacks, Hispanics, and Asians in 1990" *Social Science Research* **28** 316-326

Reardon S F, O'Sullivan D, 2004, "Measures of spatial segregation" *Sociological methodology* **34** 121-162

Schnell I, Yoav B, 2001, "The sociospatial isolation of agents in everyday life spaces as an aspect of segregation" *Annals of the Association of American Geographers* **91** 622-636

Stopher P R, 1992, "Use of an activity-based diary to collect household travel data" *Transportation* **19** 159-176

Vittengl J R, Holt C S, 1998, "A time-series diary study of mood and social interaction" *Motivation and Emotion* **22** 255-275

Wang D, Chai Y, Li F, 2011, "Built environment diversities and activity-travel behaviour variations in Beijing, China" *Journal of Transport Geography* **19** 1173-1186

Wang D, Li F, Chai Y, 2012, "Activity spaces and sociospatial segregation in Beijing" *Urban Geography* **33** 256-277

Wang D, Li F, 2015, “Daily activity space and exposure: A comparative study of Hong Kong's public and private housing residents' segregation in daily life” *Cities* (In Press) doi: 10.1016/j.cities.2015.09.010

Wong D W S, Shaw S L, 2011, "Measuring segregation: an activity space approach" *Journal of geographical systems* **13** 127-145

Table 1. A socio-demographic profile of the sample[1]

		N	Percentage	Population percentage (Census 2011[2])
Total		770	100.0	
Sex	Female	449	58.3	54.6
Age	<30	373	48.4	26.5
	30-49	283	36.8	35.5
	>= 50	114	14.8	38.0
Marital Status	Single	498	64.7	53.2
Education Level	College & above	395	51.3	37.3
Employment Status	Working	457	59.4	63.5
Monthly Household Income (HKD[3])	19999 & below	289	37.5	47.6
	20000-39999	285	37.0	29.0
	40000 & above	196	25.5	23.5
Housing Type	private housing	386	50.1	50.3
Place of Residence	Hong Kong Island	139	18.1	18.0
	Kowloon	257	33.4	29.8
	New Territories	374	48.6	52.2

[1] Several features of this table need clarification. First, the population percentages of age groups are based on population aged 10 or above. Second, since the younger group is overrepresented in the sample, and the elder group is underrepresented, the population percentages of sex, marital status, employment status, and education attainment have been weighted by age composition of the sample. Third, the population percentages of household income categories are based on household counts, while the sample percentages are based on people counts. The population percentages are thus not exactly comparable to the sample percentages, but they offer a rough estimate of how the sample matches or fails to match population characteristics.

[2] Source: Census and Statistics Department, Hong Kong SAR Government

[3] 1 HKD = 0.13 USD.

Table 2. The three dimensions of segregation in Hong Kong

Dimensions of Segregation	Residential	Work	Activity Space
Age	0.056	0.050	-0.016
Income	0.334***	0.069	0.077**
Housing	0.495***	0.129***	0.182***

(*: $p<0.01$; **: $p<0.05$; ***: $p<0.001$; similarly hereinafter.)

Table 3. Interaction between income and housing segregation

	Residential	Work	Activity Space
income	0.107	0.039	0.121
housing	0.316***	0.057	0.178***
*housing*income*	0.107	0.056	-0.088

Table 4. The moderation effect of personal mobility on housing segregation

		Residential	Work	Activity Space
Car Ownership	***housing***	**0.529*** **	**0.161*** **	**0.214*** **
	car	0.147**	0.158*	0.141**
	housing*car	**-0.169*** **	**-0.169***	**-0.160****
Size of Activity Space	***housing***	**0.564*** **	**0.148****	**0.243*** **
	sd	-0.003	-0.026	0.117**
	housing*sd	**-0.117****	**-0.033**	**-0.093**

Table 5. Housing segregation by age, gender and education attainment

		Residential	Work	Activity Space
Age	<30	0.497***	0.112	0.154***
	30-49	0.487***	0.160**	0.167***
	50+	0.447***	0.076	0.314***
Gender	Female	0.474***	0.167***	0.210***
	Male	0.524***	0.082	0.145**
Education	Secondary & below	0.474***	0.062	0.169***
	College & above	0.503***	0.164***	0.172***