

FOOD DESERTS, ACTIVITY PATTERNS, & SOCIAL EXCLUSION: THE CASE OF TOKYO, JAPAN

1. Introduction

1.1. Background and motivations

In recent years, urban policies have focused on the social phenomenon of *social exclusion*, which indicates that some people have difficulties obtaining social or personal opportunities. There are many definitions of social exclusion; therefore, it is difficult to identify one unifying attribute for all of the possible categories. Social exclusion could be influenced by poverty, poor housing, high crime environment, bad conditions for health, discrimination and racism, a low education level or family breakdown (Social Exclusion Unit, 2001; European Commission, 2002). Most previous studies examined the potential causes or predictors to define the level of social exclusion, with a main focus on individual factors (e.g., Howarth *et al.*, 1998; Burchardt *et al.*, 2002). Moreover, a number of studies defined “regular” physical and social standards from the resources of a dignified life (Schönfelder & Axhausen, 2003) with the concept of “relative” social exclusion.

More specifically, Jehoel-Gijsbers & Vrooman (2007) characterized social exclusion using the following two categories: (i) Economic/structural exclusion and (ii) Socio-cultural exclusion. In regards to the second category, which is considered in the current study, a normative standard for opportunity to access certain services can be desirable, either theoretically or empirically. That is, defined by enforcement in law or identified values and norms, the qualitative term “low accessibility” can connote a state of social exclusion regarding personal attributes or conditions.

In a general sense, accessibility has been estimated based on the distance from users to destinations (Pirie, 1979; Apparicio *et al.*, 2008; McEntee & Agyemen, 2010). Distance, or often travel time, involves cost and the potential to achieve the objective. However, with this point of view, accessibility can only be described in a physical manner, regardless of personal condition. Salomon & Mokhtarian (1998) stated, “[a]ccessibility is a measure of supply, namely potential mobility, and is not a descriptor of behavior.” Furthermore, accessibility should be considered in concert with the traveler’s constraints, such as individual socio-economic status (e.g., Kwan, 1998; Axhausen *et al.*, 2002; Páez *et al.*, 2010), health condition (Morland *et al.*, 2002; An & Sturm, 2012; Lee, 2012) or personal attitude through psychological mechanisms (Dijst *et al.*, 2008). Accessibility, therefore, can be divided into individual accessibility and locational accessibility (Dijst *et al.*, 2002). Social exclusion problems are highly associated with individual accessibility, i.e., the proximity or number of opportunities that one perceives on a personal level, which is often difficult to quantify.

1.2. Literature

Concerning accessibility to certain services, the topic of the “food desert problem” has received increased attention as a type of social exclusion, as it represents the inability of some people in certain areas to access adequate groceries. In an introduction on this topic, Wrigley (2002) stated that the term ‘food deserts’ is used to describe an area with

relatively poor access to adequate food provision. This issue is significant because it might lead to inequalities in public health (Apparicio *et al.*, 2007). In Japan, especially in deprived cities, the number of neighborhood grocery stores has declined since the 1990s (Iwama *et al.*, 2009). Furthermore, Japan has the highest ratio of elderly people (aged 65 and over), recorded as 23.1% in 2010, and the dietary situation of the elderly in Japan is deteriorating due to the decay of neighborhood stores. Because food is a basic essential for a healthy and dignified life, food-supply stores must be considered public facilities.

Many studies in this line of research have focused on area-based aggregated-level factors of food deserts. The reason lies behind the assumption that the opportunity for food access is homogeneous among areas with similar characteristics, for example, shopping environment and demographic attributes (e.g. Apparicio *et al.*, 2007; Ball *et al.*, 2009; Berg *et al.*, 2008; Bodor *et al.*, 2008; Clarke *et al.*, 2002; Huble, 2011; Sharkey *et al.*, 2009; Smoyer-Tomic *et al.*, 2006; Zenk *et al.*, 2005). Zenk *et al.* (2005) considered the poverty level and the major race in an area as food desert factors. Sharkey *et al.* (2009) analyzed the relation between proximity and density of food suppliers in an area, considering the average vehicle access, income level or employment status of households in the area. Smoyer-Tomic *et al.* (2006) explained the food desert neighborhoods as areas with higher percentage of low-income and elderly households, and relatively low proximity to supermarkets. In these examples, it is presumed that a greater percentage of people will have difficulties in areas with a higher ratio of elderly people and people with no vehicle. In contrast, studies have reported no significant association between deprivation and the availability of groceries using area-based analysis (Winkler *et al.*, 2006).

At the individual level, research has attempted to clarify which group of people is more likely to experience difficulty in shopping. These studies are based on the assumption that the degree of inconvenience, or disparity, is distinctive according to personal attributes and shopping behavior (Choi & Suzuki, 2012), which is not aggregated into zones. This assumption implies that even in an area with low food accessibility, the level of inconvenience or shopping difficulties can differ according to the personal condition. Therefore, area-based analysis could be biased. In the more recent literature, attempts were made to develop models that can explain the relationship between areal attributes and actual shopping patterns. Páez *et al.* (2010) proposed the model-based estimates of individual travelled distance to calculate accessibility. However, they focused on estimating the accessibility rather than identifying the gaps between different groups within an equal environment. These gaps may represent the meaning of social exclusion. Meanwhile, there also have been studies that investigated the risk by the loss of critical food suppliers in the region. Russell & Heidkamp (2011) emphasized the impact the loss of a major supermarket has on the current food system. Choi & Suzuki (2013) studied protection strategies for critical facilities in a food desert neighborhood.

1.3. Objectives

The objective of the present study is to develop a model to identify areas that have a large discrepancy in the probability of an individual exceeding the threshold of comfortable accessibility in grocery shopping in the city of Tokyo. A binomial level of inconvenience for grocery shopping is defined as “high” if one’s walking time for shopping exceeds a certain threshold and “low” otherwise. The probability of experiencing a high level of inconvenience is calculated using a logistic regression with the binary dependent variable while simultaneously incorporating individual constraints and aggregated

accessibility as explanatory factors. Hence, the gap between groups with different attributes can be observed in the same area. A higher gap level indicates that in that area, the inferior group has a higher probability of experiencing shopping difficulties than the other group. For a more locally focused insight, the relations between each factor are investigated by adopting a geographically weighted logistic regression (GWLR) that captures spatial variation by calibrating a regression model at different points in space. With this method, a landscape of spatially differentiated gaps of probability between groups and the influence of each factor can be depicted on a map. Overall, the main contribution of this study is that it generates the probability of each group of people experiencing shopping difficulties, which is unique within the line of food-deserts-related studies, to our knowledge. The result answers the question of which area includes people who are elderly, who do not have cars and whose life territory is limited to the area near home (unemployed) and thus have relatively higher burden in travel. The method avoids the biased assumptions that every area with a higher proportion of disadvantaged people and low accessibility is problematic, as mentioned in the literature review. Moreover, the method is examined as an effective approach to differentiate the range of calculating accessibility based on the personal condition, regarding the physical constraints. It enables the integration of the spatial opportunity for shopping and personal constraints. Ultimately, the result provides convincing evidence for urban policies to determine which areas and which groups of people require a support system. Policies can then promote the opening of new stores or protect stores in food deserts, thus improving opportunities for people to shop for groceries on a daily basis, or even can embed the local agriculture markets in appropriate places as urban food systems (Gatrell *et al.*, 2011).

2. Methodology

2.1. Accessibility and inconvenience

As Páez *et al.* (2012) described in their comprehensive review, two categories of accessibility measurement can be found in related research: a normative and a positive approach. A positive approach assigns weights for the distance from origin to the store, and sometimes for the size of the store, based on the gravity theory. A normative approach posits a critical normative value beyond which people should not be expected to travel. Numerous examples can be found in the food desert-related literature, both for the positive (e.g., Lee & Lim, 2009; Clarke *et al.*, 2002) and the normative approach. In shopping for groceries, the quality of opportunity can be determined by whether one can reach the stores that provide the desired products. The opportunity of variety is also important. The basic conception is that a person does not ‘get annoyed’ or ‘feel inconvenience’ due to the need to travel within a certain critical time to achieve his/her goal, especially in the case of walking. This is a thread of the reliability in travel time, which is often defined as the probability that the system/each traveler exceeds the predetermined threshold (Clark & Watling, 2005).

In this sense, the definition of food deserts should be differentiated for disadvantaged groups. A food desert can be defined as an area with more people in a certain group who spend a relatively higher cost for travel (e.g., time, distance), which exceeds the threshold level of inconvenience for grocery shopping.

To measure accessibility, the current study adopted a normative approach, with a

positive measurement within the normative value in terms of the reachable opportunity, which is defined as the number of grocery stores within a certain reasonable walking time and distance from the trip origin:

$$A_i^g = \sum_j I_j(d_{ij} \leq C_g), \quad (1)$$

where

i = one's spatial location

j = the location of a store j

g = a group with certain characteristics

A_i^g = accessibility of the Group g at location i

d_{ij} = distance between i and j

C_g = threshold of the convenient distance of Group g

$I_j(d_{ij} \leq C_g) = 1$ if $d_{ij} \leq C_g$, 0 otherwise.

However, there has been little agreement concerning the reasonable normative walking time. An explanation of the criteria is presented in Section 4.1.

Whereas the accessibility mentioned above is an absolute value, attempts were made to measure accessibility from a relative viewpoint. Church & Marston (2003) proposed a measure of relative accessibility for individuals with mobility challenges. Páez *et al.* (2010) proposed a deprivation indicator to define the relative accessibility deprivation and to allow for the comparison of interpersonal variations in accessibility and differences in accessibility of different types of opportunities, e.g., between retail and fast food. The concept of relative accessibility is useful for a comparative study if we assume different effects of factors on different groups of individuals, divided by the characteristics of each individual. A deprivation indicator can be made between the opportunities of elders and the reference population. In the current study, rather than relative accessibility, the concept of relative probability of inconvenience is adopted for the model. A further explanation of the model is provided in Section 4.2.

2.2. Geographically weighted logistic regression (GWLR)

Recently, the consideration of spatial autocorrelation has been increasing in the research field of the statistical analysis of geo-referenced data. The traditional regression analysis, i.e., the general linear model (GLM) and the logistic regression analysis (in the case of predicting the outcome of a categorical dependent variable) have been adopted in a number of studies. These traditional regression methods produce single regression equations to summarize the global relationships between the explanatory and dependent variables over the space. Ordinary global models cannot explain the spatially varying relationships between variables, which are known as spatial nonstationarity effects (Brunsdon, 1998). Spatial relationships are incorporated into the modeling framework through the covariance of error terms (Fotheringham *et al.*, 2002). With the geographically weighted regression model, contextually sensitive coefficients can be calculated and depict continuous distributions that form distinctive clusters, such as areas with positive or

negative values of coefficients. Hence, a judicious conjecture can be made to allow the estimates to yield local profiles that change in each cluster, reflecting socio-demographic, economic and other relevant attributes of various segments.

For the traditional logistic regression with a binary response variable, the model that predicts the dichotomous situation y_i based on a set of explanatory variables x_k can be formulated as

$$y_i^* = \frac{1}{\exp(-(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki})) + 1} = \frac{\exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki})}, \quad (2)$$

where y_i^* is the predicted value of y_i , which follows the logistic curve. In this equation, y_i^* has the boundary of 0 (as $\exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}) \rightarrow 0$) and 1 ($\exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}) \rightarrow \infty$).

To linearize the probability y_i^* to perform a regression and remove the boundary of 1/0, logit transformation is applied as the equation below:

$$Z = \ln\left(\frac{y_i^*}{(1 - y_i^*)}\right) \quad (3)$$

$y_i^*/(1 - y_i^*)$ is called the odds or likelihood ratio. It transforms Z , which is restricted to the range $[0, 1]$, into the range $[-\infty, \infty]$, while the y_i^* surface remains continuous within the range from 0 to 1. The standard linear regression model can be obtained thus:

$$Z = \beta_0 + \sum_k \beta_k x_k \quad (4)$$

Now, as an expansion, the GWLR model incorporates a set of geographical locations, taking form as the following equation:

$$Z = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} \quad (5)$$

where u_i, v_i represent the coordinates of the i -th point in space, β_k is a coefficient for the k -th explanatory variable and x_{ik} is the k -th explanatory variable.

The coefficients of the model can be expressed as

$$\beta = \begin{pmatrix} \beta_0(u_1, v_1) & \beta_1(u_1, v_1) & \dots & \beta_k(u_1, v_1) \\ \beta_0(u_2, v_2) & \beta_1(u_2, v_2) & \dots & \beta_k(u_2, v_2) \\ \dots & \dots & \ddots & \dots \\ \beta_0(u_n, v_n) & \beta_1(u_n, v_n) & \dots & \beta_k(u_n, v_n) \end{pmatrix}. \quad (6)$$

The parameter for each data point can be estimated as the following equation, considering the n by n spatial weighting matrix W .

$$\beta(i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) Y, \quad (7)$$

$$W(i) = \begin{pmatrix} w_{i1} & 0 & 0 & 0 \\ 0 & w_{i2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & w_{in} \end{pmatrix}, \quad (8)$$

where w_{ij} is the weight given to data j in the calibration of the model for data point i .

Among a variety of types of weighting functions, we adopt the Gaussian function, which is commonly used to generate the weighting for each data point. A Gaussian shape is as follows:

$$w_{ij} = \exp\left(-\frac{1}{2} \left(\frac{d_{ij}}{h}\right)^2\right),$$

where d_{ij} is a measure of the distance between data points i and j and h is a quantity known as the bandwidth.

3. Outlines of study area and data

3.1. Study area

This study focuses on the 23 wards of Tokyo. Tokyo had 8.6 million inhabitants in 2012 and this population has been increasing since 1995. The elderly population, i.e., the population aged 65 years and over, was recorded as more than 1.8 million in 2012, forming more than 21.0% of the total population. The wards that form the central Tokyo area include the central business districts, which have one of the highest job densities in the world. A population census is conducted every 5 years based on the Cho-cho-moku, the minimum unit of the Japanese census tract.

3.2. Expected factors and data used

Two main data sources are used in this study: Personal Trip (PT) survey data and the point data of retail stores.

The Personal Trip (PT) survey has been conducted 5 times in Tokyo metropolitan area since 1967. We used the latest data, collected in 2008. This survey is based on individual information. Overall, 200,000 people were sampled, of which 40,084 were in the study area, representing trips for an objective of *shopping*. Certain cases for which necessary information was omitted or that were unusual—especially those with walking time above 120 minutes, as this duration cannot be considered usual for grocery shopping—were excluded. Cases with an age below 20 years were also excluded. However, the location of origin and the trip destination were provided only in the form of the divisions of Sho-zones. Thus, the locations of individuals were assigned to the center of each Sho-zone. Tokyo's 23-ward area is an agglomeration of 255 total zones. Other categories of personal information were provided, including sex, age, possession of vehicles,

household type, type of employment, total number of trips in a day and time/mode/destination of each trip. However, each data point was recorded for one day, which may not reflect the ordinary behavior of an individual. To overcome this caveat, we excluded data outside the lower and upper bounds of 95%.

With regard to the distribution of grocery stores for calculating accessibility, the point database of retail stores was constructed from two sources. The Database of Large Retail Stores in Japan 2010 provides information on every large-scale retail store that has been recorded from 1950 to 2009. The definition of a large-scale retail store is a store that is larger than 1,000 m². We extracted only supermarkets that existed in the year 2009 because they are considered to be grocery stores. In addition, the telephone pack database in 2009 was used. This database provides information on every facility in Japan, including the locational coordinates. Facilities that are considered 'other grocery retailers' were extracted based on the codes shown in Table 1. The accessibility of supermarkets and other grocery retailers was calculated separately and considered separate variables in the estimation model.

Table 1 Extracted codes for other grocery retailers

According to the PT survey data, approximately half of the people aged 35-64 have cars (the rate is 0.500 for age 35-49; 0.523 for age 50-64). However, elderly people have a lower possession rate (0.420 for age 65-79, 0.227 for age 80+).

Table 2 describes the percentage of each travel mode with respect to car possession and age. It is clear that (i) with increasing age, more people tend to walk to stores, (ii) elderly people tend to use the bus, but not the rail or bike and (iii) most of the time, people who use a car travel less via public transportation. Thus, possession of a car can significantly influence shopping behavior.

Table 2 Travel mode share of respondents with respect to age and possession of cars

Figure 1 shows the average travel time by each mode according to age with and without cars. Concerning the average time of trips via walking, there is a significant discrepancy between the non-elderly group (age below 65) and the elderly group (age 65 and over). The average travel time of non-elderly people is less than 10 minutes, both with and without cars, whereas that of elderly people is approximately 13 minutes. This result implies that to achieve an equal level of opportunity, elderly people who walk to shopping locations must walk for more time, considering their physical restrictions, such as slower walking speed.

**Figure 1 Average travel time in minutes
(top: respondents with cars; bottom: respondents without cars)**

* Note that the share of cars is small, as presented in Table 2.

Figure 2 depicts the cumulative relative frequency of trip time via walking. Of note, the data are of 95% significance. The responses are likely to be discrete by 5 minutes, e.g., the respondents reported 5, 10, 15 or 20 minutes rather than 6, 12 or 19 minutes. This explains the exact cut value of quartiles in Table 3. Table 4 adds persuasive information that the trip time by walk can also differ according to the possession of a car and one's

employment status. Possession of a car yields larger gaps for the elderly group, whereas the gaps in non-elderly groups are slightly larger when the employment state is compared. Concerning employment status, for most age ranges, employed people spend significantly less time walking to shopping locations. This finding can be explained as employed people typically make a drop-by trip, i.e., they tend to drop by a store on the way home from their workplace.

Figure 2 Empirical cumulative distribution of trip time via walking

Table 3 Overall description of trip time via walking (in minutes)

Table 4 Comparison of walking trip time by possession of a car, employment status and age (in minutes)

4. Model frameworks

4.1. Threshold of the convenient distance

The threshold of the convenient distance can differ by physical restriction. In this study, each threshold for each group, i.e., elderly people and non-elderly people, is based on the assumption that elderly people's walking speed is slower. The average trip time of elderly people is approximately 13 minutes. Considering that the responses are discrete with 5-minute intervals, the threshold of convenient time is 15 minutes in this study. Possible arguments on threshold setting are discussed in greater depth in Section 6.1.

In calculating accessibility, the distance that one can walk with one's walking speed should be clarified. Saito *et al.* (1986) conducted an experiment on the average walking speeds of each group of Japanese elderly and non-elderly people, which are 0.68 m/s (1,210 m for half an hour) and 1.16 m/s (2,079 for half an hour), respectively. Based on this result, we set the threshold of the convenient distance for the elderly group as 600 m and that for the non-elderly group as 1,000 m.

4.2. Frame of the model

Based on the PT survey and its summary shown in Chapter 3, we constructed a logistic model that consists of a binary dependent variable and several explanatory variables, as shown in Table 5. The dependent variable was termed as *Inconvenient* because the convenience level is presumed to be "good enough" within a normative value and the individual would experience inconvenience otherwise. The benefits of this method are that both the probability of an individual exceeding the threshold convenient walking time when shopping and the probability of each group can be drawn.

Variables that represent the groups of social exclusion or food deserts were selected. As shown in Table 4, people who are elderly, do not have a car or are unemployed can be considered to more easily experience inconvenience. Each condition was set as the variables *Elderly*, *NoCar* and *Unemployed*. The variables *Elderly*, *NoCar*, *Unemployed*

were dummy coded to simplify each group and make the model more elastic, which is effective, as evidenced by the comparison results presented in Section 4.2. The dummy variable *Elderly* was 1 if an individual's age is 65 and over. The variable *NoCar* had the value of 1 if the individual's household did not possess a car. The variable *Unemployed* had the value of 1 if an individual was unemployed at the time of the survey. In addition to the individual's conditions, accessibilities were also considered, but in an adjusted form. If an individual were elderly, an accessibility of $C_{elder} = 600$ m was adopted; otherwise $C_{non-elder} = 1,000$ m was adopted. In addition, as discussed above, one's employment status influences trip time. Therefore, the accessibility of the workplace location should be supplemented. To do so, we added both accessibilities from origin and from workplace for each individual. The spatial distribution of calculated accessibilities based on equation (1) for the elderly and non-elderly groups is shown in Figure 3, aggregated by Sho-zone. The variable *ACC_Food_adj* is the adjusted accessibility of other grocery retailers by age status and the variable *ACC_SM_adj* is the adjusted accessibility of supermarkets.

Figure 3 Average accessibility of supermarkets (*ACC_SM*; top) and other grocery retailers (*ACC_Food*; bottom): Elderly group (left, $C_{elder} = 600$ m), non-elderly group (right, $C_{non-elder} = 1,000$ m)

Table 5 Description and summary of variables in the model

Before conducting the analysis, Moran's *I* is provided for each variable to verify the spatial autocorrelation in OLS residuals of the model. Moran's *I* is an index for spatial autocorrelation based on the feature locations and feature values simultaneously. Given a set of features and an associated attribute, it evaluates whether the pattern expressed is clustered, dispersed, or random.

In general, the *z*-score and *p*-value are interpreted to evaluate the significance of the index. As shown in Table 6, spatial autocorrelation is confirmed; therefore, the use of GWLR is valid.

Table 6 Moran's Index for each variable (Euclidean distance with threshold of 5 km)

5. Results

To verify the model fitness, total model fit is tested with likelihood test (Table 7) and the significance of each variable is tested using Chi-square and log-likelihood test (Table 8). Note that this test can only be conducted with the traditional OLS model. In both tests, including all variables, the model is significant.

Table 7 Full model likelihood test

Table 8 Test for significance of each variable

The GWLR was then conducted using AICc (corrected Akaike Information Criterion) minimization to determine the best-fitting model. The converged bandwidth was 14.1 km, with an AICc value of 15,071.86. Table 9 summarizes the coefficient values of

each variable.

Table 9 Parameter summaries

For binary logistic regression, another common method to verify the model fit is the classification table. Table 10 shows the classification table of the estimation samples of the dependent variable. With the cut-off value of 0.19 (calculated as $P_i^g(Y=1|X)=0.1861$, from Table 5), the overall classification power is 60.65%.

Table 10 Classification table for the estimation samples (Variable Inconvenient)

Based on equation (9), the predicted probability from the model equation can be considered the composition of individual in the same group. The full model, which includes whole variables, is expressed as follows:

$$P_i^g(Y=1) = \frac{1}{\exp(-(\beta_{0i} + \beta_{1i}Elderly_i + \beta_{2i}NoCar_i + \beta_{3i}Unemployed_i + \beta_{4i}ACC_SM_adj_i + \beta_{5i}ACC_Food_adj_i)) + 1} \quad (9)$$

Using the equation above, the probability of inconvenience for each group can be calculated. The groups can be categorized according to the combinations of dummy variables, as presented in Table 11. For example, Group 1 represents people who are elderly, do not have a car and are unemployed and thus face the most barriers or are mobility-impaired. In contrast, Group 8 represents people who are not elderly, have cars in their households and are employed. In this sense, the following expression is possible for Group 8, which substitutes the ‘pure’ effect of accessibilities only.

$$P_i^g(Y=1) = \frac{1}{\exp(-(\beta_{0i} + \beta_{4i}ACC_SM_adj_i + \beta_{5i}ACC_Food_adj_i)) + 1}$$

Table 11 Groups by combinations of dummy variables

Hence, the predicted probability of inconvenience differs by group and spatial location due to the adoption of GWLR. Figure 4 shows the average probability of each group. Group 1 has the highest probability overall. The gaps between Groups 2 and 3 and between Groups 6 and 7, which represent the standardized influence of having a car and having a workplace, are similar. The factor *Elderly* shows the largest effect in raising the probability. Figure 5 compares the cumulative relative frequency of predicted probabilities by each individual status. The gaps between the probabilities for elderly and non-elderly people are the largest, but possession of a car and employment status yield significant gaps.

Figure 4 Comparison of probability of inconvenience by groups

Figure 5 Comparison of probability of inconvenience by individual status

The predicted probabilities are summarized according to Sho-zone to determine the distribution and provide a reasonable insight. The results for Groups 1, 4 and 8 are shown in Figure 6. In every case, the central business district near Tokyo station has the highest probability. This result is assumed because the area has a significantly low population, almost no supermarkets and a high proportion of the office district. There are many more commuters than residents in this area. It is also notable that the northern-eastern area of Tokyo city has a relatively low probability, particularly in Adachi-ward and Edogawa-ward. Several reasons for this finding can be considered. For example, these areas have a high ratio of supermarkets compared to other grocery retailers (Figure 7). Moreover, specifically in the outer area of the circular railway (Yamanote line), an obvious drop in the probability of inconvenience is detected near several railway stations. It is interesting that among the stations in the central-circular line, the Ikebukuro and Shinjuku-Shibuya stations area show a distinctive gap. This gap is assumed to be caused by characteristics of those areas. As shown in Figure 6, more supermarkets are located in the Ikebukuro area (#1 in Figure 6) and very few are in the Shinjuku-Shibuya area (#2 in Figure 6).

In addition to the calculated probability, it is also meaningful to compare the areas in which a larger gap/discrepancy appears in terms of the relative probability between groups. Figure 8 shows the results of calculating the relative probability of inconvenience, with respect to that between Groups 1 and 8. In this case, the relative probability can be drawn by the ratio $p_i^1(Y=1|X)/p_i^8(Y=1|X)$. The value can be interpreted using the following example: when the relative probability is 1.4 in a certain zone in Figure 8, 40% more people are predicted to be inconvenienced in Group 1 than in Group 8 in that zone. Higher gaps appear in the eastern area, such as Edogawa-ku, Adachi-ku and Katsushika-ku. Within the central area, the Ikebukuro station district (Toshima-ku) shows a comparatively higher gap. However, a higher gap does not necessarily represent inconvenience for the mobility-impaired. For example, even in areas with high accessibility, if many shopping patrolling trips are expected, such as in the Ikebukuro station district, there may be a higher probability of inconvenience. In these cases, a longer trip time does not necessarily indicate that the person experiences inconvenience. To remove this effect, areas with low accessibility are extracted (Figure 8). For simplicity, a low-accessibility area is defined as an area in which either the accessibility of supermarkets or that of other grocery retailers within 1,000 m is below average. A high discrepancy within a low accessibility area indicates that people with barrier factors tend to feel inconvenienced due to low accessibility. As a result, the distribution of low-accessibility grocery retailers has a pattern that is more similar to a greater discrepancy than is the distribution of low-accessibility supermarkets. Nevertheless, there is an apparent difference between western and eastern districts, even within low-accessibility areas. Several districts in the eastern area have the highest discrepancies. This result provides a critical insight for the rapidly aging society because this discrepancy may increase as a greater number of elderly people struggle in the future. Moreover, it is notable that the geographical patterns of the elderly population, car possession ratio and employment rate (Figure 9) do not coincide with the areas that have a higher gap. This result provides a clear implication for the recently suggested methods to

map the food desert area.

Figure 6 Predicted probability $P_i^g (Y = 1)$ (Top: Group 1, middle: Group 4, bottom: Group 8)

Figure 7 Ratio of accessibility of supermarket to that of other grocery retailers

Figure 8 Relative probability: ratio of Group 1 to Group 8 with low accessibility

Figure 9 The percentage of elderly population (top), no possession of cars (middle) and unemployed (bottom). The ratio is calculated using only the number of respondents.

6. Discussion and conclusion

The discussion is organized as follows. We first describe possible arguments concerning the methodology and data limitations. Next, we conclude by highlighting some remaining issues and weaknesses of this study and making suggestions for future studies.

6.1. Possible arguments

The results clarify the spatial difference in the relative probability of inconvenience between groups. Specifically, the north-eastern area shows a significant gap between socially excluded and non-excluded groups. However, arguments can arise in two aspects. First, it is difficult to state that 15 minutes of walking time always yields the perception of inconvenience. For example, people with a disability or health problems can experience inconvenience or difficulty more easily, even if they are included in Group 8. Furthermore, there is an opinion that people experience greater inconvenience in shopping when they do not have sufficient spare/free time in terms of quality of life (van Eck *et al.*, 2005). That is, if an individual has a busy work schedule and grocery retailers are a far distance, the person would be quite likely to experience difficulties. These opinions premise some remarks of incorporating data on various travel patterns and lifestyles, which are difficult to collect. In this point of view, it is meaningful to detect areas with a high discrepancy and low accessibility in a macro viewpoint and draw agreement from decision makers. Furthermore, numerous applications are possible by adopting different values for the threshold.

Another expected question concerns the ambiguity of the personal trip data. Because trips vary by objective, travel mode, individual & location and even by trips that an individual makes, sophisticated data are ideal for both measuring accessibility and constructing a model. Páez *et al.* (2010) indicated that a majority of previous studies relied on the blanket-criteria-based measurement of accessibility, which is suspicious in regards to suitability and statistical reliability. However, the current study incorporates a normative concept that suggests that within a certain “desired” travel time/distance, there are no significant advantages of shorter time/distance. A more questionable problem is that the cutting method can ignore the facilities located outside the peripheral region with little distance. Fuzzy method is expected to relieve this problem, which we will utilize in a future study. In regards to the sophisticated measure of accessibility, due to the evolving

technical improvement, Ettema & Timmermans (2007) proposed a measuring skill that considers both the demographic profile of an individual and the chain of trips. This is not captured in the current study, which focuses on the specific behavior of shopping for groceries. Furthermore, the PT data used here do not necessarily represent the general behavior of each individual, though it was surveyed with a basis of the travel chain in a day. Dong *et al.* (2006) and Schönfelder & Axhausen (2003) suggested the space-temporal accessibility with a high level of quality. In fact, a common behavioral pattern gained by a repetitive observation can yield a better model fit and implications. However, it is somewhat questionable in terms of empirical application, as those tasks require high cost and time as well as a violation of private information. Considering the negative aspects, we believe that the results of this study can be adopted in the decision making process.

6.2. Remaining issues and suggestions

It is thought that the distribution of high levels of gaps is also influenced by other areal demographic factors, including population density and the percentage of elderly people. Furthermore, urban structural factors, such as the ratio of supermarkets to other grocery retailers, proportions of people who use cars or other travel modes, convenience in parking or comfortableness of the walking space, may also be influential. Market factors such as the price of groceries cannot be ignored. These factors are not included in the present study due to low significance or the unavailability of such data. However, more detailed data and a specific factor analysis are required, especially on the urban structural factors, as they directly affect people's behavior. Moreover, the areas with a high probability of exceeding the threshold regardless of relativity and the areas with a high level of gaps must be incorporated into the results.

For future studies, the definition of food deserts should be considered while accounting for the relationship between food deserts for each group. Moreover, consideration of patrolling behavior is desirable for this study. The availability of patrolling or shopping opportunities between the trip chains should be considered for greater understanding and agreement. The generalization of the probability of inconvenience is expected. Though the threshold of walking time is examined using the PT data, further research is required to clarify the actual time range that an individual does not feel inconvenience in walking, which can be investigated using a survey, for example. In light of quality of life, it is important to extend the concept of inconvenience to all methods of travel. A more in-depth consideration is needed to set the inconvenience levels for different types of trip objectives and travel modes.

Acknowledgments

This study was supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (JSPS) and Grants for Social Welfare Activities from the Mitsubishi Foundation. We appreciate their support.

References

- An, R., & Sturm, R. (2012). School and residential neighborhood food environment and diet among California youth. *American Journal of Preventive Medicine* 42 (2), 129-135.
- Apparicio, P., Cloutier, M., & Shearmur, R. (2007). The case of Montréal's missing food deserts: Evaluation of accessibility to food supermarkets. *International Journal of Health*

Geographics 6 (4), 1-13.

- Axhausen, K. W., Zimmermann, A., Schönfelder, S., Rindsfuser, G., & Haupt, T. (2002). Observing the rhythms of daily life: A six-week travel diary. *Transportation* 29 (2), 95-124.
- Ball, K., Timperio, A., & Crawford, D. (2009). Neighborhood socioeconomic inequalities in food access and affordability. *Health & Place* 15 (2), 578-585.
- Berg, N., & Murdoch, J. (2008). Access to grocery stores in. *International Journal of Behavioural and Healthcare* 1 (1), 22-37.
- Bodor, N. J., Nicholas, J., Rose, D., & Farley, T. A. (2008). Neighborhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. *Public Health Nutrition* 11(4), 413-420.
- Brunsdon, C. (1998). Exploratory spatial data analysis and local indicators of spatial association with XLISP-STAT. *The Statistician*, 47 (3), 471-484.
- Burchardt, T., Le Grand, J., & Piachaud, D. (2002). Degrees of exclusion: Developing a dynamic, multidimensional measure. In: Hills, J. Le Grand, J. and Piachaud, D. (Eds.) *Understanding Social Exclusion*. Oxford University Press, Oxford, pp. 30-43.
- Choi, Y., & Suzuki, T. (2012). A study on shopping behaviors for groceries and awareness on convenience of elderly people. *Journal of the City Planning Institute of Japan* 47 (3), 271-276 (In Japanese with English abstract).
- Choi, Y., & Suzuki, T. (2013). Protection strategies for critical retail facilities: Applying interdiction median and maximal covering problems with fortification *Journal of the Operations Research Society of Japan* 56 (1), 38-55.
- Church, R. L., & Marston, J. R. (2003). Measuring accessibility for people with a disability, *Geographical Analysis* 35 (1), 85-96.
- Clark, S., & Watling, D. (2005). Modelling network travel time reliability under stochastic demand. *Transportation Research Part B: Methodological* 39 (2), 119-140.
- Clarke, G., Eyre, H., & Guy, C. (2002). Deriving indicators of access to food retail provision in British cities: Studies of Cardiff, Leeds and Bradford. *Urban Studies* 39 (11), 2041-2060.
- Dijst, M., Farag, S., & Schwanen, T. (2008). A comparative study of attitude theory and other theoretical models for understanding travel behaviour. *Environment and Planning A*, 40 (4), 831-847.
- Dijst, M., Jong, T. d., & Eck, R. J. (2002). Opportunities for transport mode change: an exploration of a disaggregated approach. *Environment and Planning. B, Planning and Design* 29, 413-430.
- Dong, J. X., Ben-Akiba, E. M., Bowman, L. J., & Walker, L. J. (2006). Moving from trip-based to activity-based measures of accessibility. *Transportation Research Part A* 40 (2), 163-180.
- Ettema, D., & Timmermans, H. (2007). Space-time accessibility under conditions of uncertain travel times: theory and numerical simulations. *Geographical Analysis* 39 (2), 217-240.
- European Commission. (2002). *Joint Report on Social Inclusion 2001*. Office for Official Publications of the European Communities, Luxembourg.
- Fotheringham, S. A., Brunsdon, C., & Charlton, E. M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: Wiley.
- Gatrell, J. D., Reid, N., & Ross, P. (2011). Local food systems, deserts, and maps: The spatial dynamics and policy implications of food geography. *Applied Geography* 31, 1195-1196.
- Howarth, C., Kenway, P., Palmer, G., & Street, C. (1998). *Monitoring Poverty and Social Exclusion: Labour's Inheritance*. York Publishing Services, York.
- Hubley, T. A. (2011). Assessing the proximity of healthy food options and food deserts in a rural area in Maine. *Applied Geography* 31, 1224-1231.
- Iwama, N., Tanaka, K., Sasaki, M., Komaki, N., & Saito, Y. (2009). The dietary life of the elderly in local cities and food desert issues: A case study of Mito City, Ibaraki Prefecture.

- Japanese Journal of Human Geography* 61, 139-159 (In Japanese).
- Jehoel-Gijsbers, G., & Vrooman, C. (2007). *Explaining social exclusion: A theoretical model tested in the Netherlands*. The Netherlands Institute for Social Research/scp, The Hague.
- Kwan, M. (1998). Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis* 30 (3), 191-216.
- Lee, G., & Lim, H. (2009). A spatial statistical approach to identifying areas with poor access to grocery foods in the city of Buffalo, New York. *Urban Studies* 46 (7), 1299-1315.
- Lee, H. (2012). The role of local food availability in explaining obesity risk among young school-aged children. *Social Science & Medicine* 74 (8), 1193-1203.
- McEntee, J., & Agyeman, J. (2010). Towards the development of a GIS method for identifying rural food deserts: Geographic access in Vermont, USA. *Applied Geography* 30, 165-176.
- Morland, K., Wing, S., & Roux, A. (2002). The contextual effect of the local food environment on residents' diets: The atherosclerosis risk in communities study. *American Journal of Public Health* 92, 1761-1767.
- Páez, A., Gertes, R., Farber, S. M., Morency, C., & Roorda, M. (2010). Relative accessibility deprivation indicators for urban settings: Definitions and application to food deserts in Montreal. *Urban Studies* 47 (7), 1415-1438.
- Páez, A., Scott, D. M., & Catherine, M. (2012). Measuring accessibility: positive and normative implementations of various accessibility indicators. *Journal of Transport Geography* 25, 141-153.
- Pirie, G. (1979). Measuring accessibility: A review and proposal. *Environment and Planning A*, 11, 299-312.
- Preston, J., & Rajé, F. (2007). Accessibility, mobility and transport-related social exclusion. *Journal of Transport Geography* 15, 151-160.
- Russell, S. E., & Heidkamp, C., P. (2011). 'Food desertification': The loss of a major supermarket in New Haven, Connecticut. *Applied Geography* 31, 1197-1209.
- Saito, Y., Koshigaya, K., Narita, Y., Yoshida, K., & Kimura, T. (1986). A study of spastic gait on stroke patient - Comparative study of walking analysis of pre and post operation. *The Hokkaido Journal of Orthopaedics and Traumatology* 30, 115-120.
- Salomon, I., & Mokhtarian, P. (1998). What happens when mobility-inclined market segments face accessibility-enhancing policies? *Transportation Research Part D: Transport and Environment* 3 (3), 129-140.
- Schönfelder, S., & Axhausen, K. W. (2003). Activity spaces: Measures of social exclusion? *Transport Policy* 10 (4), 273-286.
- Sharkey, J., Scott, H., Daikwon, H., & Huber, J. (2009). Association between neighborhood need and spatial access to food store and fast food restaurants in neighborhoods of Colonias. *International Journal of Health Geographics* 8 (9), 1-17.
- Smoyer-Tomic, K. E., Spence, J. C., & Amrhein, C. (2006). Food deserts in the prairies? Supermarket accessibility and neighbourhood need in Edmonton, Canada. *The Professional Geographer* 58, 307-326.
- Social Exclusion Unit. (2001) *Preventing social exclusion*. Social Exclusion Unit, London.
- Van Eck, J. R., Burghouwt, G., & Dijst, M. (2005). Lifestyles, spatial configurations and quality of life in daily travel: an explorative simulation study. *Transport Geography* 13, 123-134.
- Winkler, E., Turrell, G., & Patterson, C. (2006). Does living in a disadvantaged area mean fewer opportunities to purchase fresh fruit and vegetables in the area? Findings from the Brisbane food study. *Health & Place* 12, 306-319.
- Wrigley, N. (2002). Food deserts in British cities: Policy context and research priorities. *Urban Studies* 39 (11), 2029-2040.

Zenk, N. S., Schulz, J. A., Israel, A. B., James, A. S., & Wilson, L. M. (2005). Neighborhood racial composition, neighborhood poverty, and the spatial accessibility of supermarkets in metropolitan Detroit. *American Journal of Public Health* 95 (4), 660-667.

Highlights

- The model draws a probability for an individual to exceed the threshold time.
- Food deserts are defined as areas with higher gap in probabilities.
- Areas with low food accessibility do not necessarily correspond to food deserts.
- Being elderly performs as the most influencing factor to be disadvantaged.

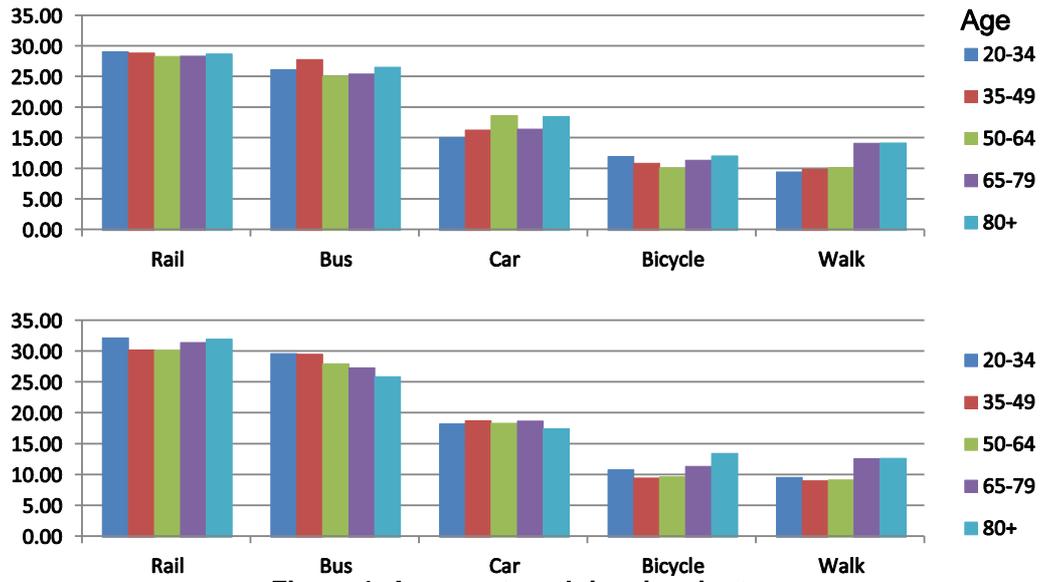


Figure 1 Average travel time in minutes
(top: respondents with cars; bottom: respondents with no cars)

* Note that the share of cars is small from Table 2.

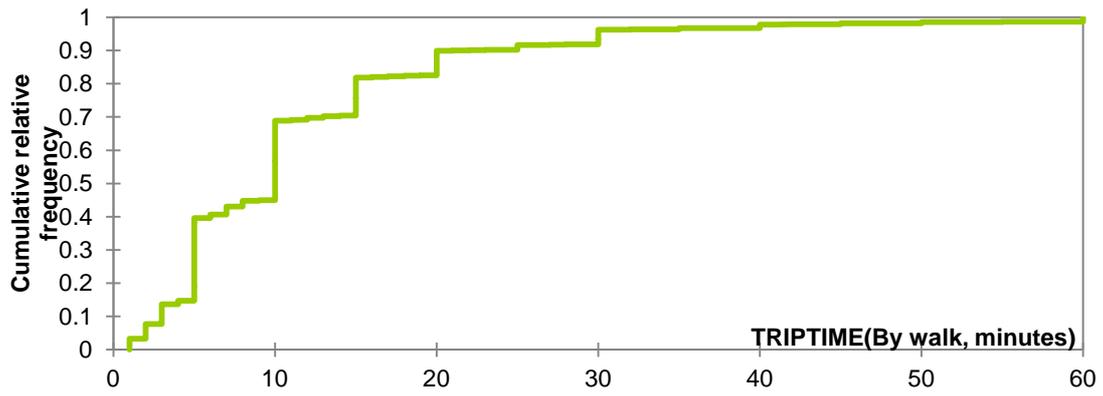


Figure 2 Empirical cumulative distribution of trip time via walking

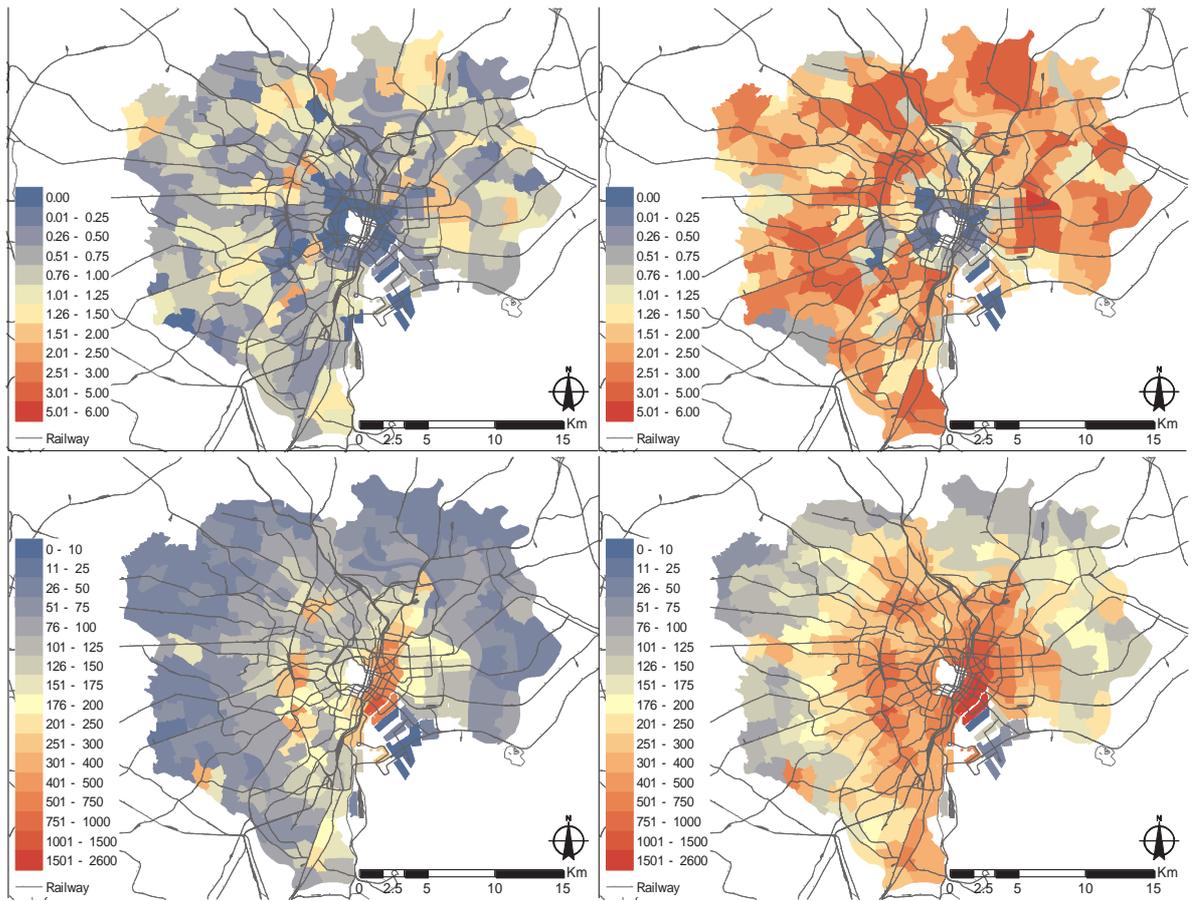


Figure 3 Average accessibility of supermarkets (ACC_{SM} ; top) and other grocery retailers (ACC_{Food} ; bottom):

Elderly group (left, $C_{elder} = 600$ m), non-elderly group (right, $C_{non-elder} = 1,000$ m)

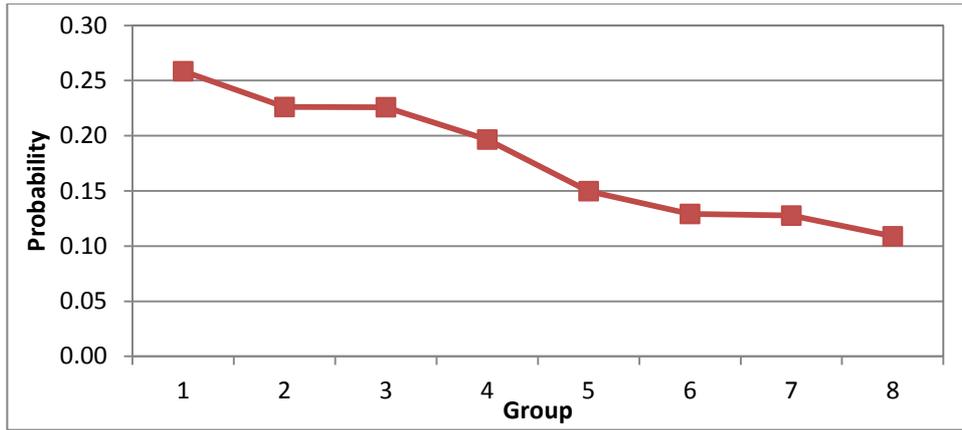


Figure 4 Comparison of probability of inconvenience by groups

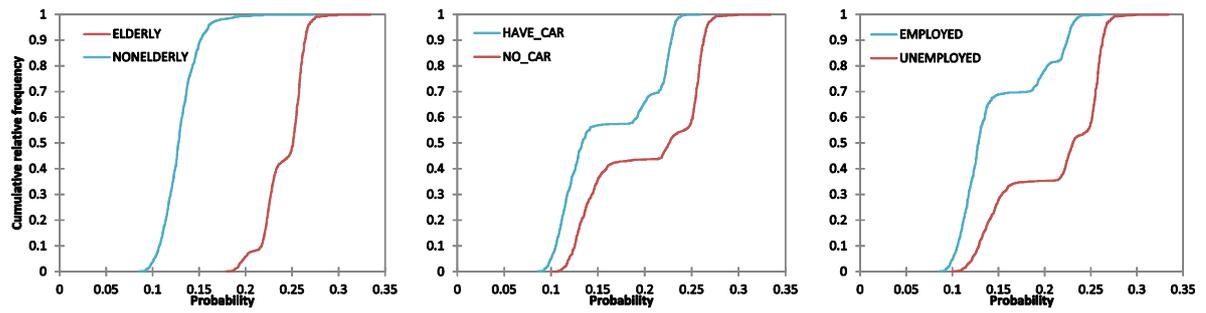


Figure 5 Comparison of probability of inconvenience by individual status

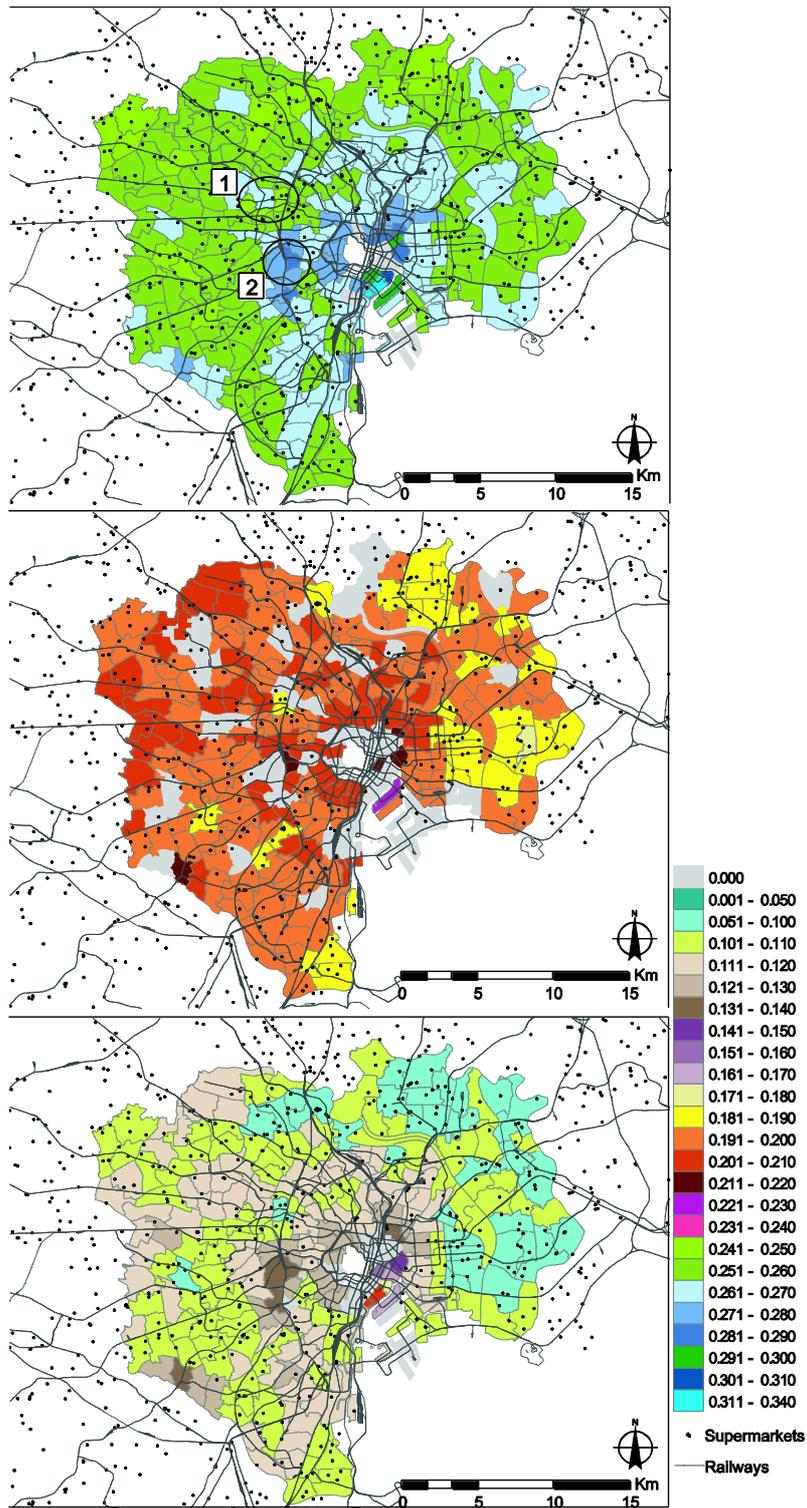


Figure 6 Predicted probability $P_i^s (Y = 1)$ (Top: Group 1, middle: Group 4, bottom: Group 8)

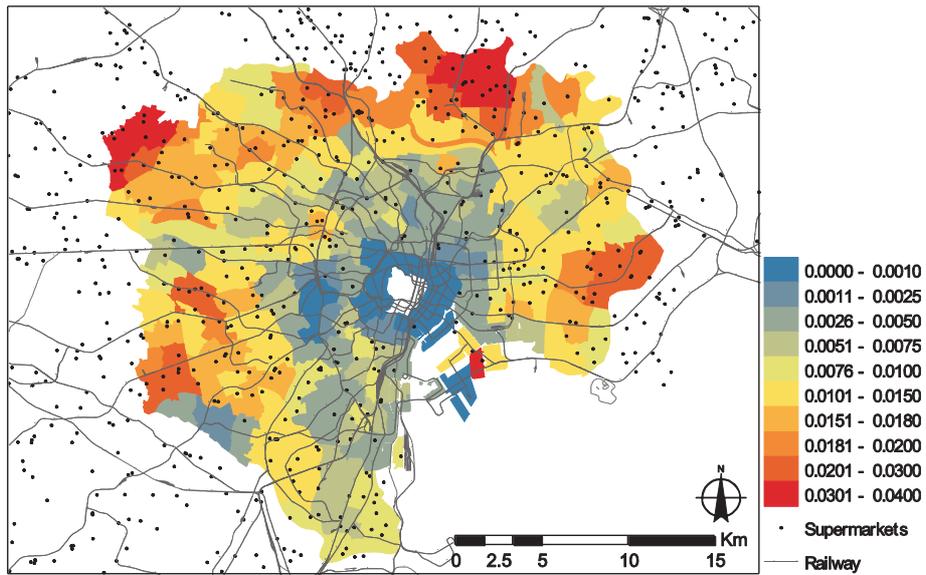


Figure 7 Ratio of accessibility of supermarket to that of other grocery retailers

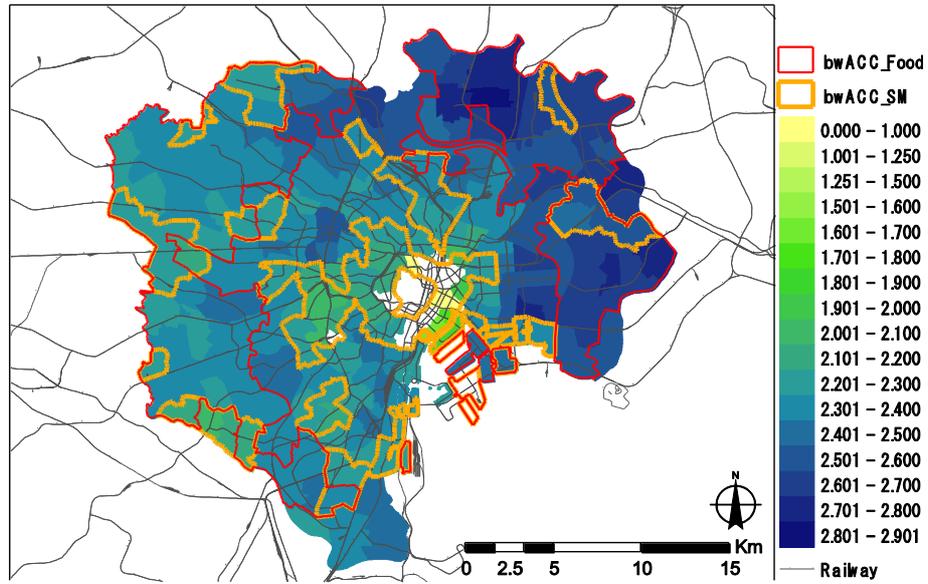


Figure 8 Relative probability: ratio of Group 1 to Group 8 with low accessibility

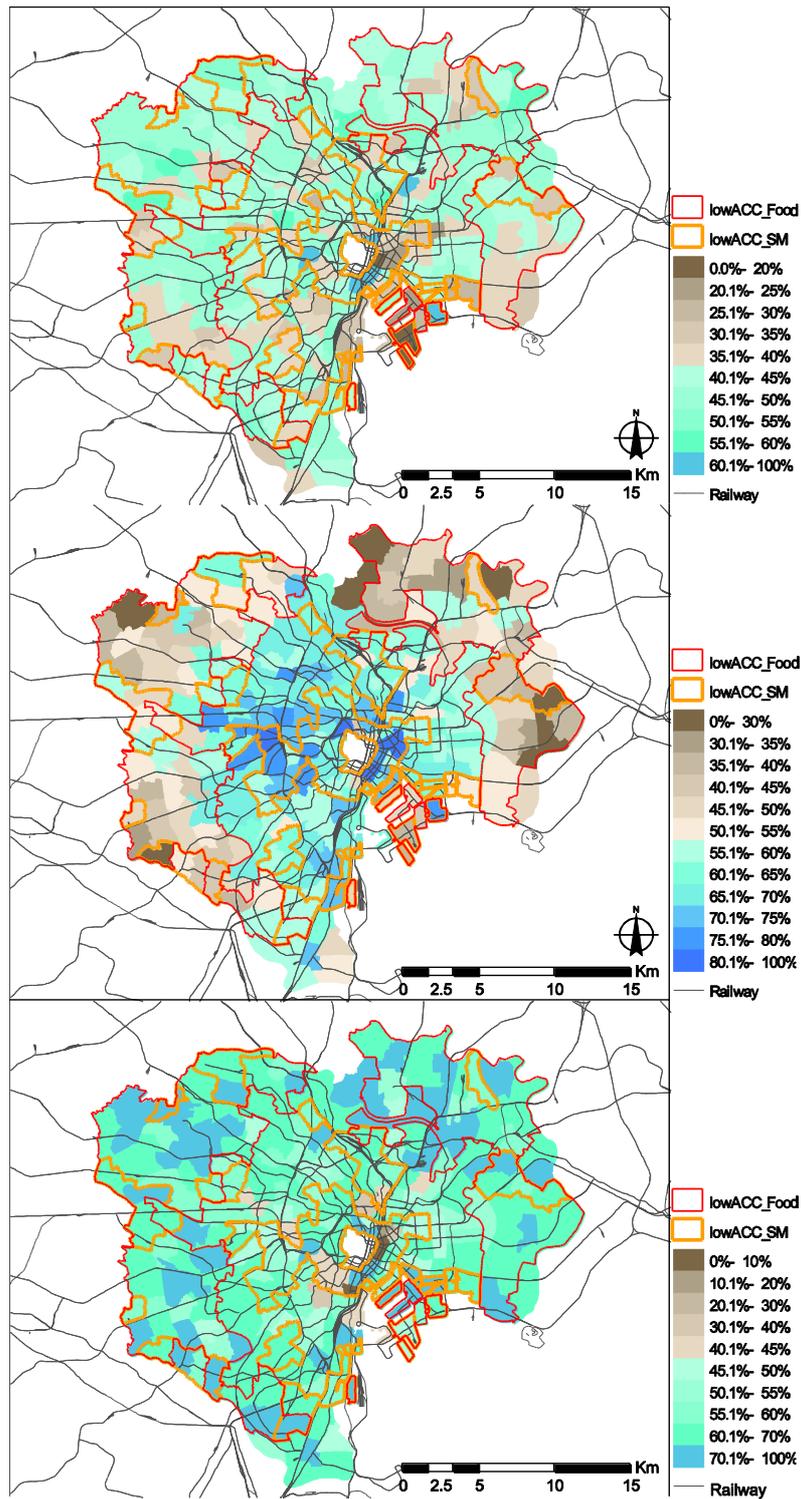


Figure 9 The percentage of elderly population (top), no possession of cars (middle) and unemployed (bottom). The ratio is calculated only by the number of respondents.

Table 1 Extracted codes for other grocery retailers

Main Category	Sub-category	Main Code	Sub-code
Groceries	Vegetables	4	10
	Livestock products	4	20
	Seafood	4	30
	Canned food (vegetables, fruits), other preserved foods	4	40
	Seasonings, herbs	4	50
	Sweeteners	4	60
	Bean jams	4	70
	Grains (rice, barley, beans, etc.)	4	80
	Oil, fat	4	90
	Salt	4	11
	Others	4	12
	Noodles	4	13
Commerce	Market places (Wholesale/fruits/fish)	19	04
Daily Life-related stores	Convenience stores	35	01
	Bakeries	35	41
	General grocery stores	35	51
	Grains	35	52
	Vegetables & Fruits	35	53
	Tea	35	54
	Milk	35	55
	Dairy products	35	56
	Meat & eggs	35	57
	Seafood	35	58
	Dried food	35	59
	Processed food	35	60

Source: Telephone pack database in 2009, ZENRIN co., Ltd.

Table 2 Travel mode share of respondents with respect to age and possession of cars

Age Car	20-34		35-49		50-64		65-79		80+	
	Yes*	No**	Yes	No	Yes	No	Yes	No	Yes	No
Rail	32.5%	39.8%	16.4%	33.6%	15.9%	26.7%	11.1%	17.2%	7.5%	11.2%
Bus	2.3%	3.1%	1.4%	2.8%	1.4%	2.8%	2.7%	5.8%	7.8%	13.9%
Car	12.3%	0.8%	16.6%	1.1%	14.9%	0.8%	17.9%	1.0%	18.1%	1.2%
Bike	18.8%	17.1%	34.0%	22.2%	35.4%	26.1%	29.6%	23.9%	15.5%	13.7%
Walk	31.5%	37.5%	30.1%	38.8%	30.3%	41.7%	35.3%	47.8%	47.6%	55.1%

(*Yes: the respondent has a car; **No: otherwise)

Table 3 Overall description of trip time via walking (in minutes)

Statistic	#Observations	Minimum	Maximum	1st Quartile	Median	3rd Quartile	Mean	Std. deviation	Skewness (Pearson)	Kurtosis (Pearson)
TRIP TIME (By walk)	16075	1	60	5.0	10.0	15.0	11.47	10.2	2.3	6.8

Table 4 Comparison of walking trip time by possession of a car, employment status and age (in minutes)

Age	Car (Yes)	Car (No)	Unemployed	Employed	Total
20-34	9.70	9.61	10.91	9.17	9.64
35-49	9.10	9.75	10.12	8.89	9.46
50-64	9.27	10.13	10.24	9.24	9.75
65-79	12.75	13.26	13.21	12.72	13.09
80+	12.80	13.72	13.53	13.55	13.53
Total	10.74	11.87	12.22	10.22	11.47

Table 5 Description and summary of variables in the model

Variable	Description	Count (value=1)	Compo- sition (%)	Average	Std. deviation
<i>Inconvenient</i>	1 if an individual's trip time exceeds 15 minutes, 0 otherwise	2,992	18.61	-	-
<i>Elderly</i>	1 if an individual's age is 65 and over, 0 otherwise	8,193	50.97	-	-
<i>NoCar</i>	1 if the individual's household does not possess a car, 0 otherwise	10,273	63.91	-	-
<i>Unemployed</i>	1 if an individual was unemployed by the time of the survey, 0 otherwise	9,952	61.91	-	-
<i>ACC_Food_adj</i>	Adjusted accessibility of other grocery retailers	-	-	205.39	203.05
<i>ACC_SM_adj</i>	Adjusted accessibility of supermarkets	-	-	1.52	1.04

Table 6 Moran's Index for each variable (Euclidean distance with threshold of 5 km)

Variable	Moran's Index	Variance	z-score	p-value
<i>ACC_SM_adj</i>	1.006	0.000002	821.64	0.0000
<i>ACC_Food_adj</i>	1.004	0.000002	819.69	0.0000
<i>Elderly</i>	0.012	0.000002	10.15	0.0000
<i>NoCar</i>	0.026	0.000002	20.99	0.0000
<i>Unemployed</i>	0.014	0.000002	11.12	0.0000
<i>Triptime</i>	0.005	0.000001	4.53	0.000006
<i>Inconvenient</i>	0.007	0.000002	6.10	0.0000

Table 7 Full model likelihood test

Statistic	DF	Chi ²	Pr > Chi ²
-2 Log(Likelihood)	5	388.849	< 0.0001

Table 8 Test for significance of each variable

Variable	DF	Chi-square (Wald)	Pr > Wald	Chi-square (LR)	Pr > LR
<i>NoCAR</i>	1	15.722	< 0.0001	15.911	< 0.0001
<i>Elderly</i>	1	94.602	< 0.0001	97.457	< 0.0001
<i>Unemployed</i>	1	14.967	0.000	15.109	0.000
<i>ACC_SM_adj</i>	1	5.228	0.022	5.265	0.022
<i>ACC_Food_adj</i>	1	5.792	0.016	5.533	0.019

Table 9 Parameter summaries

Variable	Minimum	1/4 quartile	Median	3/4 quartile	Maximum
<i>Intercept</i>	-2.1553	-2.0778	-2.0302	-2.0017	-1.9521
<i>NoCAR</i>	0.1288	0.1607	0.1748	0.1881	0.2117
<i>Elderly</i>	0.5869	0.6220	0.6555	0.6860	0.7323
<i>Unemployed</i>	0.1594	0.1744	0.1817	0.1892	0.1968
<i>ACC_SM_adj</i>	-0.0729	-0.0674	-0.0636	-0.0613	-0.0570
<i>ACC_Food_adj</i>	0.0002	0.0002	0.0003	0.0003	0.0004

Table 10 Classification table for the estimation samples (Variable Inconvenient)

	To 0	1	Total
From 0	8,111 (True Negative)	5,073 (False Positive)	61.52%
1	1,252 (False Negative)	1,639 (True Positive)	56.69%
Total	86.63%	24.42%	60.65%

Table 11 Groups by combinations of dummy variables

Group	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
<i>Elderly</i>	1	1	1	1	0	0	0	0
<i>NoCar</i>	1	1	0	0	1	1	0	0
<i>Unemployed</i>	1	0	1	0	1	0	1	0