

# Homeless Networks and Geographic Concentration: Evidence from Osaka City

Shinichiro Iwata      Koji Karato

Working Paper No. 227  
November 2007

# Homeless networks and geographic concentration: Evidence from Osaka City

Shinichiro Iwata

Faculty of Economics, University of Toyama

Koji Karato

Faculty of Economics, University of Toyama

November, 2007

## Abstract

Homeless people in Osaka City are geographically concentrated. The purpose of this paper is to examine this geographic concentration by focusing on *homeless* networks. The data we use contain information on Osaka City's homeless population by census blocks. The estimated results of a spatial autoregressive model with autoregressive disturbances show that the homeless network is significantly positive across census blocks. Networks exist in a homeless society.

**JEL classification:** C31, R23

**Key words:** Homeless networks, spatial econometrics

# 1 Introduction

The homeless in Japan are defined as people who live outdoors in tents or huts and people who sleep in cardboard boxes on the streets. Therefore, in Japan, homeless people are known as *rough sleepers* or *street people*. The numbers of these people have steadily increased in the urbanized areas of Japan since the mid 1990s and have become a major policy concern in recent years. According to the 2003 Nationwide Survey on the Actual Condition of Homeless People conducted by the Ministry of Health, Labor, and Welfare, Osaka City has the largest number of homeless people in Japan. Osaka City is the seat of the Osaka prefecture, which has the second largest economy in Japan. The Osaka City University Study Group of Urban Environmental Issues (OCUSG) (2001) has revealed the geographic distribution of the homeless in Osaka City by census block, called *cho-cho-moku*, which is the smallest Japanese administrative boundary, and found them geographically concentrated. The purpose of this paper is to examine the geographic distribution of homelessness by focusing on *homeless* networks that create geographic concentration. A homeless network, defined by Wolch, Rahimian, and Koegel (1993), comprises homeless friends, members of informal homeless communities, labor recruiters, and members of homeless support groups.

Several empirical studies have examined the spatial distribution of homelessness in metropolitan areas using intercity data (Elliott and Krivo 1991; Honig and Filer 1993; Lee, Price-Spratlen, and Kanan 2003; Park 2000; Quigley, Raphael, and Smolensky 2001). One key determinant of homelessness is the state of the housing market, as suggested by the model of O’Flaherty (1995). However, less attention has been devoted to the spatial distribution of homelessness within a city. Culhane, Lee, and Wachter (1996) used a census in New York City and Philadelphia and analyzed the previous addresses of families applying for shelter. They found that homeless families come from areas where the ratios of boarded-up housing units and the ratio of persons below the poverty level are high. Schor, Artes, and Bomfim (2003) used a census of homeless people (*street people*) in the city of São Paulo and applied a regression model to test the spatial distribution of homelessness. They found that homeless people are concentrated in built-up areas of high-rise buildings for commercial and services usage. Homeless people in São Paulo prefer these

areas, because they provide discarded materials from which to obtain income from recycling, and provide leftover food for survival. Using the 1998 Homeless Count data in Osaka City, Suzuki (2007) estimated a spatial error model. He found that homeless people, (*rough sleepers*), settle near employment agencies and a Kamagasaki *yoseba*, which is similar to the American skid row (Marr 1997; Okamoto 2007), to find a new job. A *yoseba* is located in a segregated district where labor recruiters provide jobs to day laborers. The residential area around the *yoseba* consists of inexpensive, single-room occupancy (SRO) hotels or flophouses. A large number of homeless people in Osaka City come from the Kamagasaki *yoseba*, and search for jobs around Osaka City, especially in the *yoseba*, even after becoming homeless (Aoki 2003; Mizuuchi 2003; Shima 1999). Suzuki (2007) also showed that the number of public medical care facilities and daily-needs food shops within close proximity significantly affect the spatial distribution of homelessness in Osaka City. Previous studies, however, are not concerned with homeless networks.

Similarly to Culhane, Lee, and Wachter (1996), Schor, Artes, and Bomfim (2003), and Suzuki (2007), we use intracity data provided by OCUSG (2001), rather than intercity homelessness rates data. This is because the homeless network may be very localized; consequently, intracity-level data are appropriate for estimating the location choice of the homeless. In contrast to previous studies that used intracity data, we use a spatially lagged dependent variable as well as a spatially autocorrelated error term. This model provides three important contributions to understanding geographic distribution of homelessness. First, from an econometric point of view, testing a null hypothesis of no spatial dependence on both the dependent variable and the error term are necessary when we estimate the spatial regression model (Kelejian and Prucha 1998). Culhane, Lee, and Wachter (1996), and Schor, Artes, and Bomfim (2003) did not test both of these, and Suzuki (2007) tested only a spatial autoregressive disturbance term. In this paper, we show that the null hypothesis of no spatial correlation on both the dependent variable and the error term are rejected. Therefore, we apply a spatial autoregressive model with autoregressive disturbances. Secondly, from the point of view of homeless network theory, the spatially lagged dependent variable is necessary to identify the homeless network. If homeless networks exist, homeless people obtain external benefits from the member of the network who resides in the

neighborhood. Therefore, the estimated coefficient of the spatially lagged dependent variable enables us to explore the existence of a homeless network across blocks. We find from this model that a homeless network is significantly positive across blocks. Thirdly, again from an econometric point of view, the existence of a spatial lagged dependent term changes the calculation of the marginal effects of the other control variables. These marginal effects, however, will vary with the location, because it depends on the spatial weights matrix. Following Kim, Phipps, and Anselin (2003), we apply the marginal impact of a change in a dependent variable on geographic distribution at every location on average. The average impact can be easily calculated, i.e., the coefficient is multiplied by the inverse of  $(1 - \rho)$ , where  $\rho$  is the coefficient of the spatially lagged dependent variable. Kim, Phipps, and Anselin (2003) call the inverse of  $(1 - \rho)$  a *spatial multiplier*. The spatial multiplier becomes larger than 1, because our estimated result shows that the estimated coefficient of the spatially lagged dependent variable is significantly larger than zero, but smaller than one. We find that OLS underestimates the parameter in most cases, if we do not control the spatially lagged dependent variable.

The rest of the paper is organized as follows. Section 2 presents a detailed geographic distribution pattern of homelessness in Osaka City. Characteristics of the Kamagasaki *yoseba* and networks in a homeless society are also described in this section. Section 3 introduces the theoretical model of homelessness distribution. Section 4 specifies the econometric models and discusses the identification strategies. This section also describes the data used for the econometric models and the estimated results. Section 5 summarizes the main conclusions of the paper.

## 2 Homelessness in Osaka City and Homeless Networks

Figure 1 shows the number of homeless within the census blocks in Osaka City, from the 1998 Homeless Count data (OCUSG 2001), and Figure 2 (Table 1) shows the names of areas (parks) of Osaka City used in this paper. The number of census blocks (survey points) is 1,901. To avoid double counting the OCUSG (2001) count of August 20–28, 1998, consists of two components. First, homeless people who slept in cardboard boxes, on benches, and without any form of shelter,

were counted on the nights between August 20 and 24. Secondly, those who lived in makeshift shacks of cardboard or vinyl were counted during the daytime between August 24 and 28. These counts revealed there were 8,660 homeless in Osaka City in 1998. Initially, we find that homeless people are concentrated inside the Osaka Loop Line, especially around the Midosuji Line. Why do we observe such an uneven geographic distribution pattern of homelessness (concentration rather than dispersion) in Osaka City? To explain this, we examine homeless networks.

Consider Figures 1 and 2 in more detail. First, we find that homeless people heavily cluster around Kamagasaki. Secondly, we find that homeless people are dispersed from Kamagasaki to the periphery. In these areas, we find a fairly dense concentration around the business districts of Umeda and Yodoyabashi, the shopping districts of Namba and Shinsaibashi, the electrical and electronic appliances (*den-den*) town of Nippombashi (not described in the figure), and the public parks of Osaka City, especially Osaka Castle Park (PARK 4) and Nagai Park (PARK 8).

To explain the geographic concentration in Kamagasaki, we must explain the role of Kamagasaki in the lives of homeless people. In 1999, OCUSG (2001) collected microlevel data on 672 homeless people in Osaka City. Of these, almost 60% came from Kamagasaki *yoseba*. Following Aoki (2003), Mizuuchi (2003), and Shima (1999), a *yoseba* is located in a segregated district where labor recruiters provide jobs to day laborers. Kamagasaki comprises the 11 census blocks located in the Nishinari ward, and is the biggest *yoseba* district in Japan. The residential area around the *yoseba* consists of inexpensive, SRO hotels or flophouses, called *doya* in Japanese. Labor recruiters go to the *yoseba* in the early morning, negotiate wages, workplaces etc. with day laborers, and take them to work sites. The majority of the day laborers are employed as construction workers or stevedores. On the one hand, when day laborers can get a job from labor recruiters, they obtain the daily wage, and stay at a *doya*. On the other hand, when day laborers cannot get a job, they become unemployed laborers. In this case, they collect unwanted materials, especially corrugated cardboard and aluminum cans. To exchange them for cash they carry them to a junk dealer, called a *yoseya*, located near Kamagasaki. At night, they sleep on the street, i.e., they become homeless. The severe economic recession in the second half of the 1990s and a reduction in public investment reduced the demand for day laborers. This

forced them out of the *doya* and increased the number of homeless people. In the mid 1990s, the day laborers were not homeless temporarily, but permanently. Aoki (2003) calls the gradual weakening of the day-labor market in the Kamagasaki *yoseba* as *deyosebaization*. Because large numbers of homeless people search for jobs in the *yoseba* even after becoming homeless, we find them geographically concentrated.

However, an increase in the number of homeless people increases competition among them, and consequently reduces their earnings. For this reason, many unemployed homeless who give up searching for a job in the *yoseba* leave and spread throughout the city (Mizuuchi 2003; Shima 1999). First, some locate in business districts and shopping districts because in these areas they are more likely to find a job. Secondly, some locate in the electric and electronic appliances town of Nippombashi where it is easier to collect unwanted materials. Homeless people who sleep rough in these areas occupy entrances to office buildings, sidewalks, bushes, and parking lots. Thirdly, large numbers of homeless people who live in tents or huts can be found in the parks of Osaka City. The parks are attractive for homeless people because they have free water and toilets. This geographic pattern is explained by the 1999 survey of homeless people (OCUSG 2001). The survey found that the proportion of the homeless population who had worked in the *yoseba* and were willing to work in the *yoseba* was high for those living within 1 kilometer of Kamagasaki. The proportion of the homeless population who had worked in the *yoseba* but were not willing to work in the *yoseba* was high for those living 3–6 kilometers from Kamagasaki. The proportion of the homeless population who had not worked in the *yoseba* was high for those living 7–9 kilometers from Kamagasaki. Large numbers of homeless people who have no job experience in the Kamagasaki *yoseba* are employed as low-skilled workers before they become homeless (Aoki 2003; OCUSG 2001). OCUSG (2001) shows that the greater the distance from Kamagasaki, the weaker the link to Kamagasaki.

However, previous studies have not examined the effect of homeless networks on homeless people's location, which is a central concern in this paper. The homeless network may include the following mechanisms. First, homeless friends or members of informal homeless communities share information about work and groceries. (OCUSG 2001; Wolch, Rahimian, and Koegel 1993;

Yamakita 2007). Secondly, homeless people benefit from peers, who together protect themselves against harassment from residents (OCUSG 2001; Okamoto 2007; Yamakita 2007). Thirdly, clustering of homeless people occurs because a network may provide readily available labor for labor recruiters, food services, and the availability of flophouses for homeless people (Aoki 2003; Mizuuchi 2003; Shima 1999). Lastly, homeless clustering may establish volunteer groups that support homeless people, e.g., providing a soup run (Mizuuchi 2003; Okamoto 2007; Wolch, Rahimian, and Koegel 1993). Unfortunately, we cannot directly observe the size of homeless networks percolates at the geographic level from Figure 1. Therefore, we examine the result of homeless networks empirically. Before this, we discuss the theoretical model of homeless networks.

### 3 The Theory of Homeless Networks

Assume that homeless people are all identical. Suppose each homeless person faces a choice among  $J$  alternative census blocks in the open city. Assume that homeless people who reside in  $j \in J$  benefit from not only the peers who reside in the same census block, but also the number of peers in another census block. Let  $N_j$  be the number of homeless people in  $j$ , and  $\mathbf{N}_{-j} = (N_1, N_2, \dots, N_{j-1}, N_{j+1}, \dots, N_J)$  be a vector of the number of homeless people in another census block. Then the homeless network function becomes  $H_j = H(N_j, \mathbf{N}_{-j})$ , where  $\partial H_j / \partial N_k > 0$  ( $k = 1, 2, \dots, J$ ). Furthermore, suppose that the homeless people's income is  $Y_j = Y(N_j, \mathbf{N}_{-j})$ , where  $\partial Y_j / \partial N_k < 0$ . This implies that the larger the number of homeless people in  $k$ , the lower the income in  $j$ . Finally, suppose that homeless people's utility in census block  $j$  comprises the homeless network, income, and the spatial-specific characteristic vector in  $j$  ( $\mathbf{X}_j$ ). Then, the utility function of homeless people in  $j$  can be written as:

$$V_j = V(H_j, Y_j, \mathbf{X}_j), \quad (1)$$

where  $\partial V_j / \partial H_j > 0$ , and  $\partial V_j / \partial Y_j > 0$ .

Differentiating Eq. (1), with respect to the number of homeless in  $k$ , we obtain:

$$\frac{\partial V_j}{\partial N_k} = \frac{\partial V_j}{\partial H_j} \frac{\partial H_j}{\partial N_k} + \frac{\partial V_j}{\partial Y_j} \frac{\partial Y_j}{\partial N_k}. \quad (2)$$



The first term on the RHS of (2) is assumed to be positive because of the homeless network that generates a concentration force. The second term is assumed to be negative because of competition that generates a dispersion force.

Note that the utility levels equal the fixed reservation utility level outside the city because the city is open. Then, the number of homeless people in  $j$  is determined by the following  $J$  equation system:

$$\begin{aligned}
 N_1 &= N_1(\mathbf{N}_{-1}, \mathbf{X}_1, V^*) \\
 &\vdots \\
 N_j &= N_j(\mathbf{N}_{-j}, \mathbf{X}_j, V^*) \\
 &\vdots \\
 N_J &= N_J(\mathbf{N}_{-J}, \mathbf{X}_J, V^*),
 \end{aligned}$$

where  $V^*$  is the fixed reservation utility level.

## 4 Empirical Analysis

### 4.1 Econometric Model

To examine the homeless network, in the following section, we estimate the geographic distribution of homelessness.

The point of departure is the data generating process of the number of homeless people that is assumed by the linear regression model:

$$\mathbf{N} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{3}$$

where  $\mathbf{N} = (N_1, N_2, \dots, N_J)'$  is a vector of the number of homeless people in the census block,  $\mathbf{X}$  is a matrix of geographic attributes that has a parameter vector  $\boldsymbol{\beta}$ , and  $\boldsymbol{\epsilon}$  is a random error term:  $E(\boldsymbol{\epsilon}) = 0, E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = \sigma^2\mathbf{I}$ .

Alternatively, we can start from the spatially independent model  $\mathbf{N} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ , and allow alternative specifications of the error process and spatially lagged (endogenous) variable. Specifying a first-order autoregressive error term:

$$\boldsymbol{\epsilon} = \lambda \mathbf{W}\boldsymbol{\epsilon} + \mathbf{u}, \quad (4)$$

where  $\lambda$  is the spatial autoregressive error parameter,  $\mathbf{W}$  is the spatial weights matrix based on contiguity, and  $\mathbf{u}$  is a random error term:  $E(\mathbf{u}) = 0, E(\mathbf{u}\mathbf{u}') = \sigma_u^2 \mathbf{I}$ . Specifically, the spatial weights matrix consists of positive elements  $w_{jk}$ . The elements of the spatial weights matrix are defined as:

$$w_{jk} = \begin{cases} \frac{d_{jk}}{\sum_{k \neq j} d_{jk}}, & j \neq k \\ 0, & j = k \end{cases},$$

where  $d_{jk}$  is an indicator taking the value one if the census block groups  $j$  and  $k$  share a border, and zero otherwise.

In our theoretical model, the geographic distribution of the number of homeless people is endogenously determined. To capture the spatial dependency and the homeless network (spatial externalities) in our model, we use a spatial autoregressive lagged dependent variable  $\mathbf{WN}$  (in Section 4.3 we describe  $\mathbf{WN}$  as `W_HOMELESS`). The equation for the spatially dependent model is:

$$\mathbf{N} = \rho \mathbf{WN} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (5)$$

where  $\rho$  is the autoregressive parameter for the spatial lag term. The spatial autoregressive lagged dependent variable  $\mathbf{WN}$  implies that the number of homeless people by census block unit depends on the number of homeless people in the closest-neighbors area. On the one hand, if only the competition effect exists, the sign of  $\rho$  may become negative. On the other hand, if the sign of  $\rho$  is positive, it may imply that the homeless network effect outweighs the competition effect. Therefore, a positive sign suggests the existence of networks in a homeless society.

We use the classical model specification strategy of Florax, Folmer, and Rey (2003). The classical method is a ‘specific-to-general’ or ‘bottom-up’ approach (Hendry 1979). Model specification should proceed as follows:

- Estimate the spatially independent model: Eq. (3) by means of OLS.
- Test the Lagrange multiplier test statistic  $LM_\lambda$  for  $H_0 : \lambda = 0$  and  $LM_\rho$  for  $H_0 : \rho = 0$ .

- Test the *Robust* Lagrange multiplier test statistic  $LM_\lambda^*$  for  $H_0 : \lambda = 0$  (with  $\rho \neq 0$ ) and  $LM_\rho^*$  and  $H_0 : \rho = 0$  (with  $\lambda \neq 0$ ).

The Lagrange multiplier test statistic  $LM_\lambda$  (for  $H_0 : \lambda = 0$ ) has the spatial autoregressive error model Eq. (4) as the alternative hypothesis; and  $LM_\rho$  (for  $H_0 : \rho = 0$ ) a spatial autoregressive lag model Eq. (5) as the alternative hypothesis. If both tests are insignificant, we adopt Eq. (3) as the final specification. If both tests are significant, we carry on the *Robust* Lagrange multiplier test. If  $LM_\rho^*$  is significant but  $LM_\lambda^*$  is not, we estimate Eq. (5) using maximum likelihood or spatial two-stage least squares. If  $LM_\lambda^*$  is significant but  $LM_\rho^*$  is not, we estimate Eq. (4) using maximum likelihood (Anselin 1988) or generalized method of moments for the autoregressive parameter (Kelejian and Prucha 1999). If  $LM_\rho^*$  and  $LM_\lambda^*$  are significant, using feasible spatial two-stage least squares (GS2SLS) (Kelejian and Prucha 1998), we estimate the spatial lag model (5) with a spatial error term (4) as follows:

$$\mathbf{N} = \rho \mathbf{W}\mathbf{N} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} = \lambda \mathbf{W}\boldsymbol{\epsilon} + \mathbf{u}. \quad (6)$$

Eq. (6) can be expressed compactly as a Cochrane–Orcutt-type transformation model:

$$(\mathbf{I} - \lambda \mathbf{W})\mathbf{N} = (\mathbf{I} - \lambda \mathbf{W})\mathbf{Z}\boldsymbol{\delta} + \mathbf{u}, \quad (7)$$

where  $\mathbf{Z} = (\mathbf{W}\mathbf{N}, \mathbf{X})$ , and  $\boldsymbol{\delta} = (\rho, \boldsymbol{\beta}')'$ . As a first step, we estimate the two-stage least squares estimator of  $\boldsymbol{\delta}$  in the model  $\mathbf{N} = \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}$  using the matrix of instruments that is formed as a subset of linearly independent columns of  $(\mathbf{X}, \mathbf{W}\mathbf{X}, \mathbf{W}^2\mathbf{X})$ . In the second step, the spatial autoregressive error parameter  $\lambda$  is estimated in terms of the residuals obtained via the first step and the method of moments procedure. Finally, using the moments estimators, the transformed regression model in Eq. (7) is estimated by two-stage least squares.

## 4.2 Data

The main data employed in the estimation are the 1998 Homeless Count data that we discussed in Section 2. To repeat, the number of homeless within the census blocks in Osaka City is 8,660, and the number of census blocks (survey points) is 1,901. Table 2 presents summary statistics for the census block that are used in the estimation.

In the estimation stage, we try to capture the importance of the *yoseba* to the homeless person in the following two ways. First, we include DISTANCE, the Euclidean distance between the polycentric of Kamagasaki and the polycentric of the census block, which is measured in kilometers, to capture the geographic concentration of homeless people in the Kamagasaki *yoseba* directly. We expect the number of homeless to be negatively related to the distance from Kamagasaki, because the Kamagasaki *yoseba* offers an employment opportunity, cheap food services, and SRO services for homeless people. Figure 3 shows the relationship between the number of homeless people in the census block and the distance from Kamagasaki. We find a large number of homeless at a distance of 5–6 kilometers from Kamagasaki, because Osaka Castle Park (PARK 4) and Nagai Park (PARK 8) are these areas. After controlling for these two major parks, we find that the number of homeless people decreases at a decreasing rate as we move away from Kamagasaki. Therefore we include DISTANCE SQUARED, the square of distance. The predicted sign of the square of distance is positive.

Secondly, we estimate another equation without DISTANCE and DISTANCE SQUARED, to capture the magnitude of homeless networks in the Kamagasaki *yoseba*. It is efficient for labor recruiters and the volunteer groups to find homeless people in the Kamagasaki *yoseba*, because a large number of homeless people concentrate in the Kamagasaki *yoseba*. In fact, the volunteer groups that support homeless people, e.g., providing free soup-run services, are concentrated in the Kamagasaki *yoseba* and the surrounds. This implies that homeless clustering provides an external link between the homeless individual and members of outside homeless communities. Therefore, if we control for the distance from Kamagasaki, the effect of homeless networks on the homeless population becomes weak.

Spatial-specific characteristics also include PPL, the number of nighttime residential persons who work as production process laborers in the census block. Note that production process laborers include construction workers and stevedores. As mentioned in Section 2, the majority of the day laborers who face the threat of becoming homeless are employed as construction workers or stevedores.

Because homeless people in Osaka City search for a job after becoming homeless, they will

choose to reside close to the business and commercial districts (Mizuuchi 2003, Shima 1999). Therefore we include EMPLOYEE, the number of employees in the census block. We also incorporate POPULATION, the number of nighttime residential persons (except the number of residential persons who are production process laborers) in the census block. Aoki (2003) finds that homed-family kinship networks have prevented people becoming homeless in Japan, although the ability of these networks to do so is weakening. Presently, homeless people in Osaka City are increasingly becoming separated from their homed-families and relatives. Furthermore, OCUSG (2001) finds that 20% of homeless people suffer from harassment by residents. Therefore, homeless people are less likely to choose to reside close to residential areas. This implies that POPULATION should be negatively related to the number of homeless.

We control for STATION, the number of train stations within 500 meters of the census block, because homeless people live in stations (OCUSG 2001). Transportation facilities operating in Osaka City include Japan Railway, private railways, and subways. There are four private railway companies: Hankyu; Hanshin; Kintetsu; Nankai; and seven subway lines and one new tramline operated by the Osaka Municipal Transportation Bureau. We include all stations on these lines that operate in Osaka City. Following Suzuki (2007), we make WELFARE FACILITY, the number of welfare facilities within 500 meters of the census block. Welfare facilities offer a minimum standard of living, i.e., social and health care services. We also control for the size of the census blocks (SIZE), because the census blocks vary in size and shape. Finally, we include 12 parks as dummy variables (PARK 1–12) in Osaka City (see Figure 2), because homeless people live in parks (OCUSG 2001).

In previous studies, the ratio of affordable (low-quality) housing is important because homeless people come from such areas. If homeless people are likely to settle near their previous addresses, we have to control for it. However, we do not directly control this variable, because DISTANCE and PPL may be positively correlated to it. As mentioned, Kamagasaki *yoseba* consists of inexpensive SRO hotels or flophouses. Production process laborers may dwell in low-rent neighborhoods, although the data are not available. Therefore, DISTANCE and PPL may proxy the low quality housing.

### 4.3 Estimation Results

Table 3 provides the OLS estimation results for Eq. (3) and the Lagrange multiplier tests based on the OLS residuals. The model in column (b) adds DISTANCE and DISTANCE SQUARED to column (a).

From column (a), we find that all estimated coefficients are significant, except the census block size (SIZE). The number of homeless people increases with respect to the nighttime population of production process laborers (PPL). As expected, homeless people settle near the business and commercial districts where there are many employees (EMPLOYEES), but stay away from residential suburbs where there is a large nighttime population (POPULATION). PPL has the largest magnitude among three variables, which are measured in units of population. This result may reflect that a large number of homeless persons are employed as production process laborers before they become homeless, and locate in the same neighborhoods even after becoming homeless. The result also implies that homeless people strongly prefer to locate in low-rent neighborhoods, because PPL may proxy the low-quality housing, as mentioned in Section 4.2. We also find that homeless people occupy train stations (STATION) and live near welfare facilities (WELFARE FACILITY). The coefficients of park dummies (PARK 1–12) have a statistically significant positive effect on the homeless population. They also have similar values as shown in Table 1, except for PARK 12.

From column (b), the number of homeless people decreases at a decreasing rate the further the distance from Kamagasaki. This implies that homeless people are geographically concentrated in Kamagasaki and the surrounding areas. The size of the census block (SIZE) is significant at the 10% level. More space increases the homeless population.

Important for our argument is the Lagrange multiplier test statistics  $LM_\rho$  and  $LM_\rho^*$ , which have a Chi-squared distribution with one degree of freedom. All null hypotheses that the autoregressive parameter is zero are rejected, because the  $p$ -values (square brackets in Table 3) are sufficiently small. These imply that the spatial lag term (W\_HOMELESS), i.e., the proxy of the homeless network, must be considered. Furthermore, the Lagrange multiplier test statistics  $LM_\lambda$  and  $LM_\lambda^*$ , which have a Chi-squared distribution with one degree of freedom, reject the

null hypothesis. Therefore, model Eq. (6) is estimated.

Table 4 shows the estimated results of Eq. (6) by means of GS2SLS. Again, the model in column (b) adds DISTANCE and DISTANCE SQUARED to column (a). The hypothesis that the spatial autoregressive error is not present ( $\lambda = 0$ ) is rejected at the 10% significance level in column (a), and the 5% significance level in column (b), respectively. After considering the spatial autoregressive error, the coefficient for the number of homeless people in the closest-neighbors area (W\_HOMELESS) has a statistically significant positive impact on the number of homeless people. This has implications for the geographic concentration of homeless people by census block unit; as the number of homeless people in the closest-neighbors area increases, the number of homeless people in the area increases. This suggests that a homeless network exists in a homeless society.

As expected, the coefficient of W\_HOMELESS in column (a) is larger than that in column (b). The effect of homeless networks becomes strong, because the distance from the Kamagasaki *yoseba*, where homeless networks exist, is not controlled for in column (a). From column (b), we again find that geographic concentration in Kamagasaki, because DISTANCE and DISTANCE SQUARED are statistically significant with the expected signs, even if we control for homeless networks.

Let us move to the other control variables. To interpret the magnitude of the other control variables, we have to consider a spatially lagged dependent term, because it changes the calculation of the marginal effects of the other control variables. Following Kim, Phipps, and Anselin (2003), we obtain:

$$J^{-1} \sum_k \sum_j \frac{\partial N_k}{\partial X_{ij}} = \frac{\beta_i}{(1 - \rho)}. \quad (8)$$

The parameter  $\beta_i(1 - \rho)^{-1}$  measures the effect on the homeless population of a unit change in the  $i$ -th dependent variable ( $X_i$ ) at every location on average, except the distance from the Kamagasaki *yoseba*. The spatial multiplier ( $(1 - \rho)^{-1}$ ) becomes larger than one, because our estimated results show that both  $\rho$  in columns (a) and (b) are significantly larger than zero, but less than one. Substituting estimated values in columns (a) and (b) of Table 4 into Eq. (8), we obtain the average impact. Comparing these values with the marginal effects (coefficients) in

Table 3, we find that OLS underestimates in most cases. From this point of view, the spatially lagged dependent variable corrects for the bias of OLS.

Finally, we compare the coefficients of the distance to Kamagasaki (DISTANCE and DISTANCE SQUARED) in column (b) of Table 3 with that in Table 4. Define  $D_j$  as distance from Kamagasaki of the  $j$ -th observed point. On the one hand, from Eq. (3), the marginal effect of mean distance from Kamagasaki on average is:

$$J^{-1} \sum_k \frac{\partial \hat{N}_k}{\partial D_j} = \hat{\beta}_1 + 2\hat{\beta}_2 J^{-1} \mathbf{D}' \mathbf{i}, \quad (9)$$

where  $\hat{N}_k$  is the fitted value,  $\hat{\beta}_1$  ( $\hat{\beta}_2$ ) is the estimated values for DISTANCE (DISTANCE SQUARED),  $\mathbf{D} = (D_1, D_2, \dots, D_J)'$  is a vector of the Euclidean distance between the polycentric of Kamagasaki and the polycentric of the census block,  $\mathbf{i}$  is the vector, which contains a column of ones. Substituting estimated values in column (b) of Table 3 into Eq. (9), we obtain  $-5.49$ . On the other hand, from Eq. (6), the marginal effect of mean distance from Kamagasaki on average is:

$$J^{-1} \sum_k \sum_j \frac{\partial \tilde{N}_k}{\partial D_j} = \frac{\tilde{\beta}_1}{1 - \tilde{\rho}} + 2\tilde{\beta}_2 J^{-1} \mathbf{D}' ((\mathbf{I} - \tilde{\rho} \mathbf{W})^{-1})' \mathbf{i}, \quad (10)$$

where  $\tilde{N}_k$  is the fitted value,  $\tilde{\beta}_1$  ( $\tilde{\beta}_2$ ) is the estimated values for DISTANCE (DISTANCE SQUARED), and  $\tilde{\rho}$  is the estimated values for W\_HOMELESS. Substituting estimated values in column (b) of Table 4 into Eq. (10), we obtain  $-7.91$ . Therefore the differential coefficient in column (b) of Table 3 is larger than that in column (b) of Table 4. This implies that OLS underestimates the geographic concentration in Kamagasaki, compared with the GS2SLS method.

## 5 Conclusion

The purpose of this paper is to examine homeless networks in geographic areas in Osaka City. We use the 1998 homelessness data for Osaka City from OCUSG (2001), which provides homeless population by census block, and we estimate a spatial autoregressive model with autoregressive disturbances. This model has three important contributions to understanding the geographic distribution of homelessness. First, from an econometric point of view, testing a null hypothesis



of no spatial dependence on both the dependent variable and the error term is necessary when we estimate the spatial regression model (Kelejian and Prucha 1998). In this paper, we show that the null hypothesis of no spatial correlation on both the dependent variable and the error term is rejected. Therefore, we apply a spatial autoregressive model with autoregressive disturbances. Secondly, again from an econometric point of view, the OLS estimator is biased when a spatial lagged dependent term is significant. We find the presence of a downward bias in OLS estimates. Thirdly, from the point of view of homeless network theory, this model enables us to explore the existence of homeless networks across census blocks. We find that the number of homeless people in a census block is significantly influenced by the number of homeless in neighboring census blocks, because the spatial lagged dependent term is significantly positive. This suggests that homeless networks are one factor in homelessness concentration by census block units.

The Kamagasaki *yoseba* (day-labor market) has a strong influence on the geographic distribution of homelessness in Osaka City. First, for the case where we include the distance from Kamagasaki, we find that homeless people are geographically concentrated in Kamagasaki and the surrounding areas. This is because, the Kamagasaki *yoseba* offers an employment opportunity, cheap food services, and SRO services for homeless people. Secondly, for the case where we exclude the distance from Kamagasaki, we find that the magnitude of homeless networks becomes strong. This implies that homeless networks do exist in the Kamagasaki *yoseba*.

From the estimated results, we also find the previous address is important in explaining the distribution of the homeless population. A large number of homeless people were employed as production process laborers in the Kamagasaki *yoseba* (day-labor market) before they become homeless. Therefore, not only the distance from Kamagasaki, but also the population of production process laborers in census blocks are statistically significant in the expected way. This result also implies that homeless people strongly prefer to locate in low-rent neighborhoods, because production process laborers may live in these areas.

Our empirical results criticize the city government policy that has been implemented since the late 1990s. The city government has evicted and dispersed homeless people, because neighboring residents and business people suffer a negative externality from the presence of homelessness.

This dispersion policy, however, threatens the life of homeless people further, because they benefit from homeless networks that create geographic concentration.

## **Acknowledgment**

The authors would like to thank Masahiro Ashiya, Mototsugu Fukushige, Tatsuo Hatta, Nobuhiro Hosoe, Tatsuhiko Kawashima, Kazuyuki Nakamura, Toru Mizuuchi, Ryosuke Okamoto, Fumio Takuma, Komei Sasaki, Takatoshi Tabuchi, Piyush Tiwari, Hiroyuki Yamada, Mitsuyoshi Yanagihara and seminar participants at Nagoya University, GRIPS, the JEA conference in Fukushima, AsRES/AREUEA conference in Vancouver, and ARSC conference in Hiroshima for their valuable comments. The authors would also like to thank Toru Mizuuchi, Yuji Ookura, and Wataru Suzuki for the data used. All remaining errors are the sole responsibility of the authors. Part of this paper was written while Shinichiro Iwata was visiting CIRJE, University of Tokyo. He is grateful for its hospitality. This research was supported by MEXT.KAKENHI (17730159).

## References

- Anselin L (1988) *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Boston
- Aoki H (2003) Homelessness in Osaka: globalization, yoseba and disemployment. *Urban Studies* 40: 361–378
- Culhane DP, Lee CM, Wachter SM (1996) Where the homeless come from: a study of the prior address distribution of families admitted to public shelters in New York City and Philadelphia. *Housing Policy Debate* 7: 327–365
- Elliott M, Krivo LJ (1991) Structural determinants of homelessness in the United States. *Social Problems* 38: 113–131
- Florax RJGM, Folmer H, Rey SJ (2003) Specification searches in spatial econometrics: the relevance of Hendry’s methodology. *Regional Science and Urban Economics* 33: 557–579
- Hendry DF (1979) Predictive failure and econometric modelling in macro-economics: the transactions demand for money. In: Ormerod P (ed) *Economic Modelling*. Heinemann, London
- Honig M, Filer RK (1993) Causes of intercity variation in homelessness. *American Economic Review* 83 (1993) 248–255
- Kelejian HH, Prucha IR (1998) A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* 17: 99–121
- Kelejian HH, Prucha IR (1999) A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review* 40: 509–533
- Kim CW, Phipps TT, Anselin L (2003) Measuring the benefits of air quality improvement: a spatial hedonic approach, *Journal of Environmental Economics and Management* 45: 24–39

- Lee BA, Price-Spratlen T, Kanan JW (2003) Determinants of homelessness in metropolitan areas. *Journal of Urban Affairs* 25: 335–355
- Marr MD (1997) Maintaining autonomy: The plight of the Japanese yoseba and the American skid row. *Journal of Social Distress and the Homeless* 6: 229–250
- Mizuuchi M (2003) Growth of rough sleepers in Osaka and the recent evolution of actions of government, NPO and volunteer organizations. In: Nakagawa S, Sumrongthong B (eds) *What's Happening on the Street?* UCRC Bangkok Office, Bangkok
- O'Flaherty B (1995) An economic theory of homelessness and housing. *Journal of Housing Economics* 4: 13–49
- Okamoto Y (2007) A comparative study of homelessness in the United Kingdom and Japan, *Journal of Social Issues* 63: 525–542
- Osaka City University Study Group of Urban Environmental Issues (2001) *Report of General Survey Concerning Homelessness*. [In Japanese]
- Park J (2000) Increased homelessness and low rent housing vacancy rates. *Journal of Housing Economics* 9: 76–103
- Quigley JM, Raphael S, Smolensky E (2001) Homeless in America, homeless in California. *Review of Economics and Statistics* 83: 37–51
- Shima K (1999) *The Homeless in Contemporary Japan*. Gakubunsha, Tokyo [In Japanese]
- Suzuki W (2007) What determines the spatial distribution of homeless people in Japan? *Applied Economics Letters* [Forthcoming]
- Wolch JR, Rahimian A, Koegel P (1993) Daily and periodic mobility patterns of urban homeless. *Professional Geographer* 45: 159–169
- Schor SM, Artes R, Bomfim VC (2003) Determinants of spatial distribution of street people in the city of São Paulo. *Urban Affairs Review* 38: 592–602

Yamakita T (2007) Comradeship within communities of homeless people. *Japanese Sociological Review* 57: 582–597 [In Japanese]

Fig.1. Homeless in Osaka City.

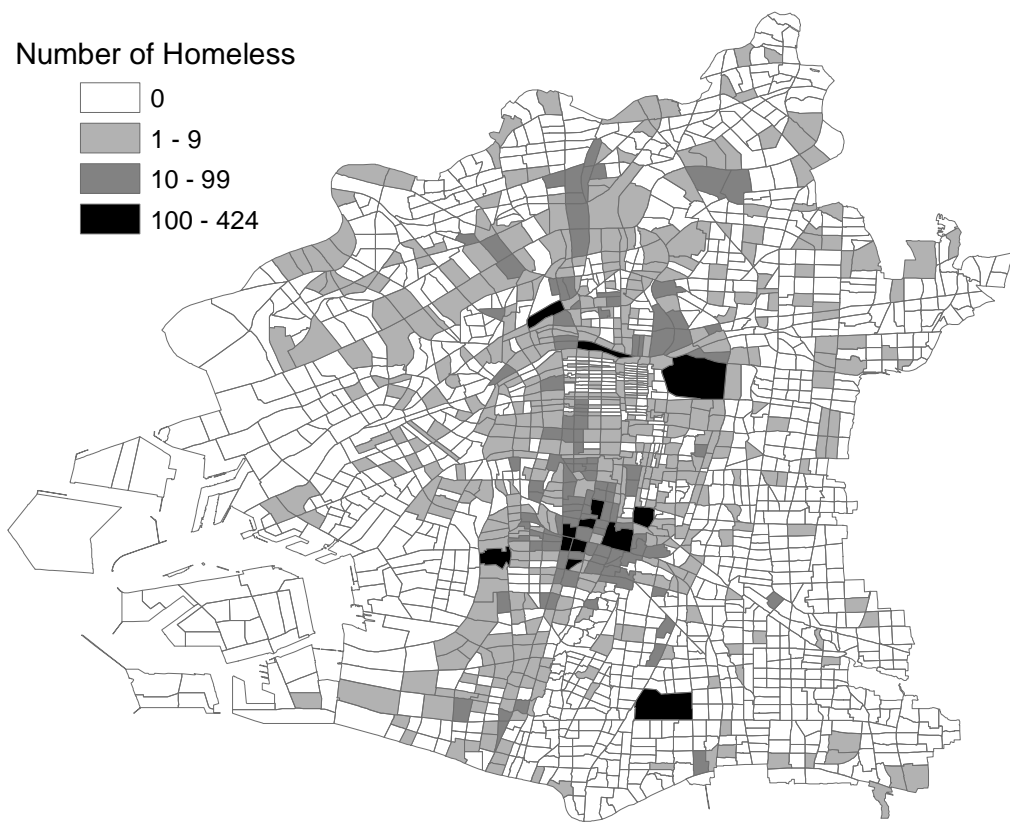
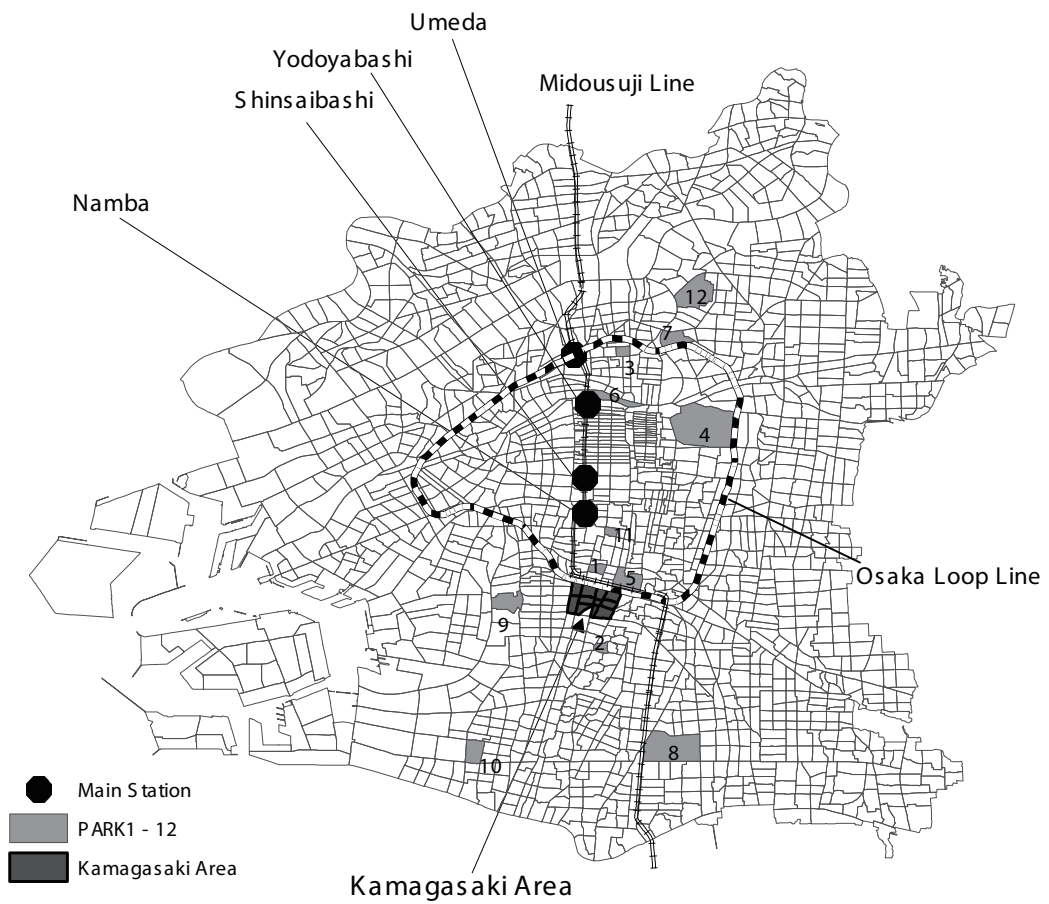


Fig.2. Names of Area of Osaka City.



**Fig.3. Homeless and the Distance from Kamagasaki.**

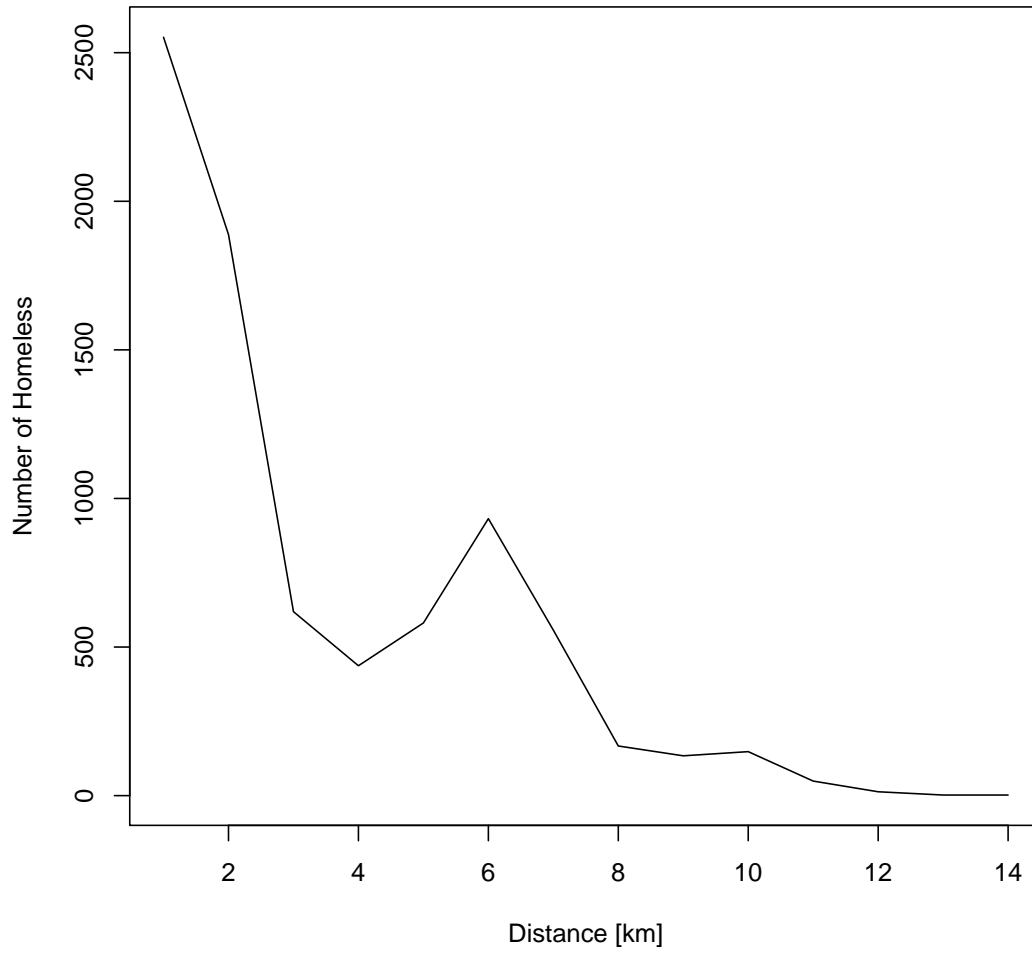




Table 1

Homelessness in main park

	Park name	Homeless (people)
PARK 1	Emi	171
PARK 2	Shotenyama	51
PARK 3	Ogimachi	95
PARK 4	Osaka Castle	350
PARK 5	Tennoji	424
PARK 6	Nakanoshima	120
PARK 7	Sakuranomiya	14
PARK 8	Nagai	310
PARK 9	Nishinari	253
PARK 10	Suminoe	25
PARK 11	Nippombashi	86
PARK 12	Kema Sakuranomiya	9

Table 2

Descriptive statistics of variables in a census block

Variable	Mean	SD	Min	Max	Source
HOMELESS (people)	4.25	23.65	0.00	424.00	HCD
DISTANCE (kilometers)	5.49	2.71	0.11	14.03	GIS
PPL (1000 people)	0.15	0.15	0.00	2.65	PC
EMPLOYEES (1000 people)	1.28	1.96	0.00	26.33	EEC
POPULATION (1000 people)	1.37	1.12	0.00	17.62	PC
STATION (number)	1.04	1.24	0.00	9.00	GIS
WELFARE FACILITY	0.19	0.44	0.00	2.00	GIS
SIZE (ha)	11.71	12.93	0.54	291.55	GIS
PARK 1	-	-	0	1	GIS
PARK 2	-	-	0	1	GIS
PARK 3	-	-	0	1	GIS
PARK 4	-	-	0	1	GIS
PARK 5	-	-	0	1	GIS
PARK 6	-	-	0	1	GIS
PARK 7	-	-	0	1	GIS
PARK 8	-	-	0	1	GIS
PARK 9	-	-	0	1	GIS
PARK 10	-	-	0	1	GIS
PARK 11	-	-	0	1	GIS
PARK 12	-	-	0	1	GIS

HCD: 1998 Homeless Count data.

EEC: 2001 Establishment and Enterprise Census.

PC: 2000 Population Census.

Table 3

## OLS and Lagrange Multiplier tests

Variable	(a)		(b)	
	Coef.	<i>t</i> -statistic	Coef.	<i>t</i> -statistic
CONST.	-3.006***	-2.99	15.084***	9.38
DISTANCE			-5.796***	-11.91
DISTANCE SQUARED			0.378***	10.00
PPL	86.370**	2.29	87.277***	20.76
EMPLOYEE	0.962**	2.43	1.421***	7.62
POPULATION	-8.883**	-2.28	-8.935***	-15.09
STATION	3.605***	4.74	2.314***	7.20
WELFARE FACILITY	2.146**	2.16	1.336*	1.73
SIZE	0.013	0.69	0.051*	1.84
PARK 1	158.274***	69.87	150.373***	10.54
PARK 2	55.306***	36.25	43.469***	3.05
PARK 3	89.281***	63.32	95.898***	6.74
PARK 4	356.624***	77.23	353.490***	24.35
PARK 5	425.598***	217.25	411.944***	28.89
PARK 6	115.990***	88.09	116.548***	8.20
PARK 7	28.733***	3.75	32.703**	2.29
PARK 8	314.870***	101.84	311.089***	21.70
PARK 9	246.303***	127.60	237.359***	16.69
PARK 10	20.787***	21.62	23.143	1.63
PARK 11	80.371***	119.09	72.758***	5.12
PARK 12	80.688**	2.49	82.343***	5.21
Adj. R <sup>2</sup>	0.608		0.640	
Number of Obs.	1901		1901	
Moran's I	18.24	[.000]	14.58	[.000]
$LM_\lambda$	324.89	[.000]	203.66	[.000]
$LM_\rho$	325.49	[.000]	187.85	[.000]
Robust $LM_\lambda^*$	71.75	[.000]	52.00	[.000]
Robust $LM_\rho^*$	72.36	[.000]	36.19	[.000]

\*\*\* indicates significant at 1%. \*\* indicates significant at 5%. \* indicates significant at 10%.  
*p*-value in square brackets.

Table 4

GS2SLS

Variable	(a)		(b)	
	Coef.	<i>t</i> -statistic	Coef.	<i>t</i> -statistic
CONST.	-3.066***	-3.91	8.970***	4.71
W_HOMELESS	0.304***	11.79	0.239***	8.60
DISTANCE			-3.812***	-6.47
DISTANCE SQUARED			0.247***	5.51
PPL	76.771***	17.92	78.840***	18.51
EMPLOYEE	0.968***	5.12	1.241***	6.50
POPULATION	-7.865***	-13.10	-8.058***	-13.54
STATION	2.084***	6.31	1.562***	4.66
WELFARE FACILITY	1.817**	2.25	1.344*	1.68
Size	0.023	0.84	0.044	1.58
PARK 1	133.321***	9.76	130.886***	9.71
PARK 2	54.723***	4.02	48.852***	3.63
PARK 3	93.100***	6.83	95.728***	7.12
PARK 4	362.264***	26.01	359.553***	26.13
PARK 5	416.359***	30.62	409.678***	30.46
PARK 6	116.265***	8.56	116.065***	8.66
PARK 7	27.352**	2.00	29.347**	2.18
PARK 8	320.984***	23.38	317.351***	23.40
PARK 9	251.696***	18.48	246.860***	18.34
PARK 10	23.549*	1.73	24.291*	1.81
PARK 11	67.451***	4.95	64.469***	4.79
PARK 12	70.307***	4.60	72.371***	4.80
$\lambda$	0.150*	1.67	0.156**	2.06
Adj.R <sup>2</sup>	0.645		0.653	
Number of Obs.	1901		1901	

\*\*\* indicates significant at 1%. \*\* indicates significant at 5%. \* indicates significant at 10%.