Spatial Analysis of Residential Burglaries in London, Ontario

Jacek Malczewski[1], Anneliese Poetz[2] and Luigi Iannuzzi[3]
[1&3]Department of Geography, University of Western Ontario, London, Ontario, Canada N6A 5C2
[2]School of Geography and Geology, McMaster University, Hamilton, Ontario Canada L8S 4K1

This paper focuses on analyzing the spatial pattern of residential burglaries in London, Ontario. It discusses the problems associated with using geo-referenced data on residential burglary incidents. The relative risk ratio is applied as a measure of the intensity of residential burglaries. The highest relative risks of residential burglaries are found in the core area of the city and the risks tend to decline with increasing distance from the city center. Another distinctive feature of the spatial pattern of residential burglaries is the west-east division. The east section of the city is characterized by higher risks of residential burglaries. The relationships between the spatial pattern and the contextual neighborhood variables are analyzed using the standard (global) multiple regression and the geographically weighted regression. It is demonstrated that the spatial pattern of residential burglaries is significantly related to the spatial patterns of socio-economic characteristics. Also, there are significant local variations in the relationships between the relative risk of residential burglary and the socio-economic characteristics of neighborhoods in London, Ontario.

Keywords: Spatial pattern of residential burglaries; global and local regression analysis, London, Ontario.

Researchers in criminology are faced with numerous theories and sets of propositions, often in conflict with one another, that attempt to explain the nature of certain empirical relationships (Bursik and Grasmick 1993). Early perspectives on criminological theory and research were guided by the assumption that the explanation for crime can be found within the individual (Davidson 1981). The traditional emphasis of offenders has been displaced in recent years by analyses of the environments in which crime occurs (Brantingham and Brantingham 1991; Ratcliffe and McCullagh 1999; Anselin et al. 2000; Craglia et al. 2000). The theory of social disorganization is a prominent component of the environmental perspectives on crime (Bursik 1988; Bursik and Grasmick 1993). It grew out of the research conducted by the Chicago school of human ecology (Burgess 1929; Shaw and McKay 1972), which documented the relationships between urban structure and the distribution of crime. The basic premises were that crime had a spatial order; it was not distributed randomly but showed trends to cluster in poorer, disadvantaged environments. Subsequent studies have shown a remarkable persistence of these basic qualities of the spatial pattern of crime (e.g. Kohfeld and Sprague 1988; Ceccato et al. 2002). The results indicate that macro-level indicators of "concentrated disadvantage" are among the strongest and most stable predictors of crime across empirical studies (Davidson 1981; Bursik 1993). It should be noted that many studies on the relationships between the socio-economic characteristics of neighborhoods and the spatial distribution of crime have been based on the victims' data. Specifically, there is a large volume of literature on the risks of victimization and the socio-economic characteristics of neighborhoods (Davidson 1981; Smith and Jarjoura 1989; Brantingham and Brantingham 1991; Ratcliffe and McCullagh 1999; Ceccato et al. 2002). Those studies have generally shown stronger relationships between social class and crime. Housing costs affect the population composition of a neighborhood, which, in turn, influences the distribution of social groups with differential risks for offending (Davidson 1981; Kohfeld and Sprague 1988; Bursik and Grasmick 1993; Bowers and Hirschfield 1999).

Crime mapping and spatial analysis of crime are essential elements of the macro-level research in criminology (Block et al. 1995; LaVigne and Wartell 1998; Messner et al. 1998; Weisburd and McEwen 1998; Harries 1999; Anselin et al. 2000). The most important aspects of spatial crime analysis involve two related tasks: (i) identifying areas of the high crime concentration or "hot spots" (Messner et al. 1998; Ratcliffe and McCullagh 1999; Anselin et al. 2000; Craglia et al. 2000) and (ii) exploring the relationship between the spatial pattern of crime and socio-economic characteristics (Bowers and Hirschfield 1999; Ceccato et al. 2002). This paper focuses on these two aspects of spatial crime analysis. It analyses the spatial variation of residential burglaries in London, Ontario and attempts to find the relationship between the spatial pattern of residential burglaries and socio-economic characteristics of neighborhoods.

This study is based on a data set consisting of 8534 residential burglary incidents during the time period from 1998-2001 in London, Ontario. The burglary incidents datasets (point data) are integrated with the socio-economic datasets (polygon/area data). The areal unit of observation is a census tract (CT) for the 1996 Population Census (Statistics Canada 1996). The analysis includes 71 census tracts in London, Ontario. The data are stored, manipulated, and analyzed using ArcView GIS 3.2 (Ormsby and Alvi 1999). In addition, multiple regression software is used for analyzing the global and local relationships between the relative risk of burglary and socio-economic characteristics of the census tracts (Insightful Corporation 2001; Fotheringham et al. 2002).

While the conventional (or global) multiple regression modeling has proved useful in research attempting to explain the variations in crime incidents (Ceccato et al. 2002), it is believed that this type of modeling adopts an unrealistic approach to an analysis of the "neighborhood effect" (Anselin et al. 2000). It is argued that the results of the multiple linear regression hide local variations in the relationships of burglary incidents and the neighborhood's socio-economic indicators (Fotheringham et al. 2002). Consequently, the geographically weighted regression (GWR) (Brunsdon *et al*. 1996; Fotheringham et al. 2002) has been applied to examine the local variations in the relationships between relative risks for residential burglaries and the socio-economic characteristics.

Data

This study is based on two sources of data. The first data source is the Statistics Canada's 1996 Census of Population (Statistics

Table 1: Summary of raw and corrected data collected from the London Police for burglaries in 1998-2001 for London, Ontario.

Number of Incidents	Year				Total
	1998	1999	2000	2001	
London Police records	1926	2400	2301	1907	8534
Deleted Records	1	10	12	17	40
%	0.05	0.42	0.52	0.89	0.47

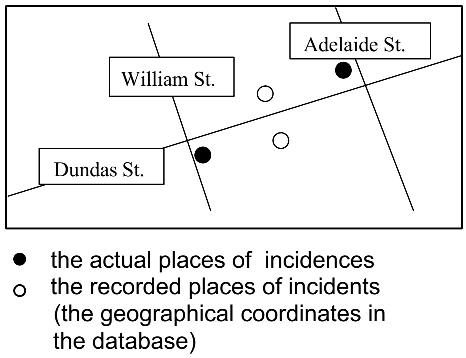

Figure1: Hypothetical street map depicting the type of data problems incurred by inaccurate geographical coordinates for point data.

Canada 1996). The Census data provide information about socio-economic characteristics of the study area by census tract (CT).

The analysis is confined to the 71 "urban" census tracts in the City of London, Ontario. The second set of data is a digital data file for residential break-ins. The datasets, provided by the London Police, consist of four comma-delimited files; one file for each year 1998, 1999, 2000, and 2001. Each file contains: street address, type of crime, date, and geographical coordinates. The number of incidents recorded by the London Police Department for each year is shown in Table 1. Only break-and-entry data are used in this study. Specifically, data on either damage to or theft from a vehicle that was parked at the residence or charges of domestic violence are excluded from this research. Using ArcView 3.2, we plotted a point pattern of the records based on the latitude and longitude coordinates. Then, a spatial overlay of polygon coverage of CTs was used to identify a burglary records that were plotted within the "urban" area of the city.

There are two problems associated with using the point data for analyzing spatial pattern of burglaries. First, the geographical coordinates do not correspond to the "true" locations of the break and enter incidences (see Figure 1). According to the police's reporting practice, the geographical coordinates for burglary occurrences mark (approximately) the centre of the block face, which are recorded using the first address at an intersection on the street and the last address (at the next intersection). For example, for a (hypothetical) break and enter at 605 Dundas St., the coordinate for this incidence would fall at 635 Dundas St. The address at this halfway point is determined by estimating the mid-point between the address at Dundas and Adelaide (601 Dundas St.), and the address at Dundas and William (670 Dundas St.). The rationale for using the centre of block faces is to allow dispatchers to guide the officers to the address as quickly as possible (Swalwell 2002).

Second, in some instances, the geographical coordinates for the address are missed. The explanation by the police was that there are some block face records that may not be updated in their data set due to new streets, extensions to current streets etc. (Swalwell 2002). By inputting the whole address including street number, name, and type of street, the batch-match function in ArcView 3.2 was used (ESRI 1996). In cases where the unit number was given in addition to the street address (e.g. for apartments within a multi-family building) the unit number had to be deleted. Where the name included a "direction" in its name such as North Centre Road, the entry had to be edited such that ArcView 3.2 would not misinterpret the "north" in the name for the direction. In some instances, the street name had to be edited so that the last word in the name was properly categorized as the "street type". For example, if the *street type* was "Close" or "Green", ArcView 3.2 included it as part of the *street name* field automatically, rendering it a "no match". Once this was corrected the addresses were "matches" and capable of being plotted. Still, there were problems with certain addresses for unknown reasons. The total number of address that could not be matched was 40 (see Table 1). A total of 8494 burglary incidences were included into the analysis. The incidents were geographically aggregated using the 71 CTs. The total number of burglaries for 1998-2001 in each CT was divided by 4 to obtain an average number of burglaries per year. These average values were the basis for analyzing the spatial pattern of residential burglaries.

Spatial Pattern of Relative Risks for Residential Burglaries

An exploratory mapping of rates or proportions is a first step in any spatial crime analysis (Hirschfield et al. 1995; Messner et al. 1998; Anselin et al. 2000; Ceccato et al. 2002). One simple approach is to map the raw observed rates. The rate is calculated by dividing the number of observed incidents in a given area by its "population", and multiplying by 100. However, since variability of the rates is a function of the "population" to which they relate this approach may be misleading. There are a number of alternatives that take into account the spatial variability of the underlying "population at risk" (Bailey and Gatrell, 1995; Ceccato et al. 2002; Haining 2003). In this study, the intensity of residential burglaries is measured by the relative risk (*RR*) ratio, which is defined as follows (Bailey and Gatrell, 1995):

$$RR_i = \frac{b_i}{\mu_i}100 \text{ ,where } \mu_i \ =$$

is the expected number of burglaries in the *i*-th census track, b_i is the average observed number of residential burglaries in 1998-2001 in the *i*-th census tract, and u_i is the "population at risk" (the number of dwellings) in the *i*-th census tract. Thus, the RR_i ratio is obtained by dividing the observed by the expected number of burglary incidents and multiplying by 100. If the RR_i values are less than 100, then a census tract is characterized by relatively low risk (that is, the risk is less than expected for the study area). The RR_i values greater than 100, this indicates that the risk is greater than expected. Figure 2 shows the spatial pattern of the RR_i ratios. The Jenks optimization procedure was used for classifying the ratios (see the map legend). The procedure ensures the internal homogeneity within classes while maintaining the heterogeneity among classes (Jenks and Caspall 1971).

The spatial pattern of the relative risk ratios can be examined in the context of the classical Burgess' zonal model (Burgess 1929) and the Hoyt's sectoral model of the city (Hoyt 1939). To this end, a concentric ring pattern and a sector pattern were superimposed on the RR_i ratio's map (Figure 2). The zonal model has three concentric rings (core area, middle zone, and peripheral zone) and the sectoral model consists of four sectors: north-west (NW), south-west (SW), south-east (SE), and north-east (NE). The lines dividing the city into the four sectors correspond roughly with Dundas St. and Adelaide St. The former divides the city into the north and south portions of London and the latter constitutes the boundary between the west and east sectors of the study area.

The highest RR_i values concentrate in the core area of the city and the lowest relative risk ratios are found in the peripheral areas. The geographical distribution of the relative risks, with a concentration of crime in the central part of the city, that progressively decreases as we move toward the peripheral areas, suggests a pattern similar to the classical Burgess model of the city (Burgess 1929; Clark 1982). This observation is supported by the literature, which has concluded on several occasions that the central business district is an area of high crime and the intensity of crime tends to decrease with the increasing distance from the city center (e.g. Bowers and Hirschfield, 1999; Kohfeld and Sprague, 1988).

Another distinctive feature of the spatial pattern of residential burglaries in London is the west-east division. In general, the relative risks of residential burglary tend to be higher in the eastern section of the city then in the west. This observation suggests that the pattern of crime in London can be described using the notion of the "dual city", which identifies a polarization of the city into two distinct spatial units according to socio-economic status (Marcuse 1993; Van Kempen 1994). In London, one can identify the contrast between the west of Adelaide areas of low relative risks/high socio-economic status and the east of Adelaide areas of high risks/low socio-economic status. Disadvantaged neighbourhoods are more attractive to burglars, despite the common belief that affluent neighbourhoods are at greater risk (Wiles and Costello 2000; Ceccato et al. 2002). According to Wiles and Costello (2000), "residential areas with high offence and victimization rates are generally found on poorer housing areas". This is because even poor areas contain plenty of suitable targets. In addition, "offenders tend to live in these areas and, on the whole, tend to offend close to home rather than conduct long-range instrumental searches across a city".

Although the notion of the "dual city" provides an insight into the geographical distribution of relative risks for residential burglaries in London, the spatial pattern is more complex than the east-west division. In addition to the concentration of crime in the core area of the city and the contrast between the east and west of Adelaide, there are considerable variations in relative risks of residential burglaries within sectors. This observation is supported by the literature. For example, Marcuse (1993) challenges the notion of the "dual city", arguing that the divided city is not necessarily dual, but is divided into "quarters" or sectors. Figure 2 shows that the SW sector is characterized by the lowest relative risks of residential burglaries in all zones, while the SE

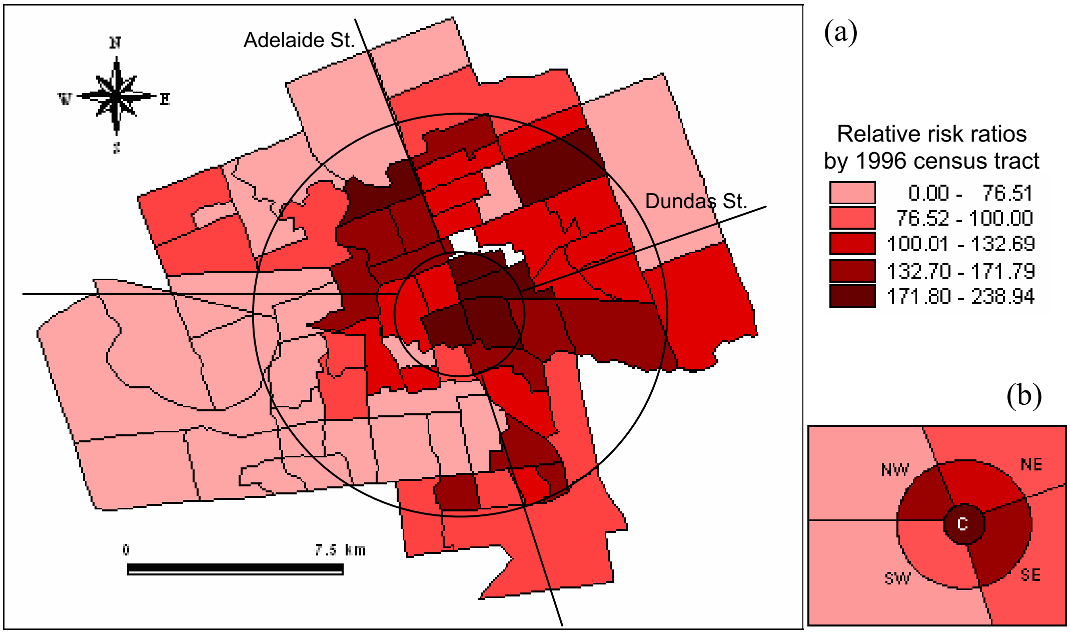

Figure 2. Map of relative risks of residential burglaries in London, Ontario, with superimposed concentric zones and sectors (C = core area; NW= north-west; NE = north-east; SE = south-east; SW = south-west).

sector of the city contains the highest RR_i values. These two sectors are also different in their socio-economic status; SW is relatively affluent, while the SE sector is characterized by low socio-economic status in all its zones. This observation suggests that the spatial pattern of crime in London corresponds to the sectoral model of the city (Hoyt 1939; Clark 1982). It should be noted however that the sectoral model allows for variation within sectors. The NW sector - containing the highest-income neighbourhoods in London - is characterized by the greatest variations in the relative risks of residential burglaries. While the peripheral zone of the NW sector is a low risk area, there are "pockets" of very high risk for residential burglaries within the middle zone of the sector. These high-risk areas are located in the vicinity of the University of Western Ontario campus, characterized by large transient-student populations. Previous studies have found that rental properties with transient populations are at higher risk for break-ins (Bowers and Hirschfield 1999).

Residential Burglaries and Socio-Economic Structure

Global Multiple Regression Analysis

In order to examine the relationships between the relative risk for residential burglaries and the socio-economic characteristics of neighborhoods, we calibrated the global multiple linear regression model (see e.g. Rogerson 2001). The relative risk (RR_i) ratios constituted the dependent variable in the model. Twelve socio-economic characteristics were included into the analysis as the potential explanatory (independent) variables (see Table 2). The choice of independent variables was based on a review of the relevant literature (Poetz 2003). Tables 2 and 3 provide descriptive statistics of all the variables and the Pearson's correlation coefficients for the dependent variable (the relative risk ratio) and the twelve independent variables. Most of the independent variables show correlations in the direction which

Table 2: Socio-economic variables considered for inclusion in the multiple regression model

Socio-economic variables	Symbol	Minimum	Maximum	Mean	Std. Deviation
Percentage of not attending school (population 15 to 24 years)	x_1	0.949	19.828	7.183	3.985
Percentage of population with less than a grade 9 education (polulation 15 years and over)	x_2	7.895	62.500	31.311	10.908
Percentage of university educated (population 15 years and over)	x_3	7.713	65.203	27.406	14.683
Percentage of visible minority population (population total)	x_4	0.986	22.189	8.688	5.256
Percentage of population who have immigrated since 1961 (total population)	x_5	12.719	33.373	20.746	5.101
Average value of dwelling in $10,000's	x_6	6.520	27.440	14.627	3.763
Median household income in $10,000's	x_7	2.546	9.127	5.130	1.366
Unemployment rate (population 15 years and over)	x_8	2.500	18.500	9.903	3.600
Percentage of government transfer payments (population 15 years and over)	x_9	4.400	30.400	15.076	6.334
Percentage of incidence of low income (population 15 years and over)	x_{10}	3.300	44.200	19.800	9.712
Percentage of population without income (population 15 years and over)	x_{11}	0.000	10.610	5.845	2.211
Percentage of non-movers within the last year (population total)	x_{12}	60.271	93.333	80.791	7.390

most observers would expect. Also, most of the independent variables show significant relationships with the relative risk ratio (see Table 3). It should be pointed out however that these correlations are likely to be affected by the inter-correlations with other variables. Until the inter-correlations have been controlled for, it is unreasonable to assume that any of the socio-economic variables has a causal connection with the residential burglary ratio, and the apparent relationship may be an artifact of relationships with a third variable. The usual way to control for such relationships is multiple regression analysis.

Given the twelve independent variables under consideration, one can generate $2^{12} - 1 = 4095$ regression models. There are number of methods for selecting the best model (Miller 1990; Selvin 1998). We use the Mallow's Cp procedure for selecting the "best" subset of the independent variables to be included into the multiple regression model. The Mallow's Cp procedure evaluates all possible models based on the mean squared error (see Miller 1990). The minimum value of the Cp measure indicates the "best" subset of independent variables that is both relatively unbiased and a good representation of the data. We used the S-Plus *leaps*( ) function to select the "best" multiple regression model in terms of the Mallow's Cp criterion (Selvin 1998; Insightful Corporation 2001). Table 4 lists the "best" models for a given number of independent variables. The minimum value of the Mallow's $Cp = 5.092$. Accordingly, the "best" multiple regression model includes the following independent variables: x_2, x_4, x_9, x_{10}, x_{11}, and x_{12}. The parameters of the model were estimated using ordinary least squares method (see e.g. Selvin 1998; Rogerson 2001). The results are given in Table 5. The model explains 64.5% of the variance in the relative risks. All the

Table 3: The Pearson's correlation coefficients for the dependent (*RR*) and independent variables (see Table 2 for the variables' description)

	RR	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
RR	1.000	.508**	0.389**	-0.269*	-0.064	-0.052	-0.313**	-0.388**	0.531**	0.395**	0.519**	0.149	-0.162
X_1	0.508**	1.000	0.667**	-0.725**	-0.006	0.217	-0.661**	-0.664**	0.619**	0.780**	0.561**	0.119	-0.125
X_2	0.389**	0.667**	1.000	-0.619**	-0.049	0.034	-0.733**	-0.768**	0.637**	0.678**	0.650**	-0.102	-0.439**
X_3	-0.269*	-0.725**	-0.619**	1.000	-0.149	-0.207	0.744**	0.582**	-0.534**	-0.619**	-0.345**	-0.413**	-0.073
X_4	-0.064	-0.006	-0.049	-0.149	1.000	0.749**	-0.027	-0.181	0.189	0.029	0.265*	0.577**	-0.124
X_5	-0.052	0.217	0.034	-0.207	.749**	1.000	-0.074	-0.208	0.142	0.152	0.230	0.418**	-0.047
X_6	-0.313**	-0.661**	-0.733**	0.744**	-0.027	-0.074	1.000	0.768**	-0.517**	-0.640**	-0.539**	-0.034	0.253*
X_7	-0.388**	-0.664**	-0.768**	0.582**	-0.181	-0.208	0.768**	1.000	-0.751**	-0.831**	-0.855**	0.020	0.636**
X_8	0.531**	0.619**	0.637**	-0.534**	0.189	0.142	-0.517**	-0.751**	1.000	0.762**	0.798**	0.263*	-0.484**
X_9	0.395**	0.780**	0.678**	-0.619**	0.029	0.152	-0.640**	-0.831**	0.762**	1.000	0.749**	-0.024	-0.393**
X_{10}	0.519**	0.561**	0.650**	-0.345**	0.265*	0.230	-0.539**	-0.855**	0.798**	0.749**	1.000	0.013	-0.742**
X_{11}	0.149	0.119	-0.102	-0.413**	0.577**	0.418**	-0.034	0.020	0.263*	-0.024	0.013	1.000	0.270*
X_{12}	-0.162	-0.125	-0.439**	-0.073	-0.124	-0.047	0.253*	0.636**	-0.484**	-0.393**	-0.742**	0.270*	1.000

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 4: The results of the S-Plus *(leaps)* procedure for selecting the independent variables for multiple regression model

Variables included in the regression models	Mallow's *Cp*
X_8	25.224
X_2, X_8	20.497
X_2, X_3, X_8	17.139
X_2, X_4, X_{10}, X_{11}	10.835
$X_2, X_5, X_9, X_{10}, X_{12}$	5.720
$X_2, X_4, X_9, X_{10}, X_{11}, X_{12}$	5.092
$X_2, X_4, X_5, X_9, X_{10}, X_{12}, X_{11}$	5.360
$X_2, X_3, X_4, X_5, X_9, X_{10}, X_{11}, X_{12}$	6.179
$X_2, X_3, X_4, X_5, X_8, X_9, X_{10}, X_{11}, X_{12}$	7.852
$X_1, X_2, X_3, X_4, X_5, X_8, X_9, X_{10}, X_{11}, X_{12}$	9.454
$X_1, X_2, X_3, X_4, X_5, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}$	11.000
$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}$	13.000

Table 5: Global multiple regression model for relative risks of residential burglary in London, Ontario.

Independent Variables	Coefficients Value	Std. Error	*t*value	Significance Pr (>I t I)
(Intercept)	-113.163	81.372	-1.391	0.169
x_2	7.000	1.493	4.689	0.000
x_4	-3.092	0.962	-3.214	0.002
x_9	-3.933	1.011	-3.890	0.000
x_{10}	4.324	1.020	4.238	0.000
x_{11}	1.828	0.952	1.951	0.049
x_{12}	3.552	2.286	1.554	0.125

regression coefficients except the one associated with x_{12} are statistically significant at the 5% level. The coefficient associated with x_{12} is statistically significant at the 12.5% level.

The variables x_2, x_{10}, x_{11}, and x_{12} all have a positive relationship with the dependent variable. Thus, the model suggests that other things being equal, the higher the percentage of population with less than a grade 9 education in a census tract, the higher the relative risk of residential burglary in that area. This finding is supported by other studies such as Reilly and Witt (1992), Timbrell (1990), and Kohfeld and Sprague (1998). There are also positive relationships between the percentages of population with a low income and without income and the relative risk of residential burglary. These relationships suggest that high relative risks of residential burglary occur in the more economically deprived areas (Bowers and Hirschfield 1999; Paternoster and Bushway, 2001; Ceccato *et al.*, 2002). The model suggests that the higher the proportion of non-movers in a given area, the higher the relative risk of residential burglary. This finding contradicts some previous studies that have found that transient populations are at higher risk for break-ins (e.g. Bowers and Hirschfield, 1999). However, Ceccato *et al.* (2002) have found that areas characterized by a lower rate of turnover were a greater target for residential burglaries.

The variables x_4 and x_9 have negative relationships with the dependent variable. The higher is the proportion of visible minorities, the lower is the relative risk for residential burglary. The negative relationship found in our study contradicts the findings of other case studies, especially in the United States (Sampson and Groves 1989; Smith and Jarjoura 1989). Also, the results from the regression model show that the higher is the proportion of the population that receives government transfer payments (welfare), the lower is the relative risk of residential burglary. It appears that this finding contradicts the positive relationship between the economically disadvantage populations and high relative risks of residential burglary. However, low-income earners may include recently unemployed persons who are receiving unemployment insurance. Unlike welfare payments, the government considers unemployment insurance as "taxable income". Others included in the low-income statistics (aged 15 years and over) could encompass teenage runaways, teenage single parents, young families, and the working-poor, who may live in multi-family public housing (Bowers and Hirschfield, 1999; Paternoster and Bushway, 2001). This suggests that the low-income segment of population may not overlap with the population receiving government transfer payments. Therefore, the relationship between the percentage of population on welfare and the relative risk of residential burglary is different than that between the low-income variable and the level of relative risk of residential burglary.

Geographically Weighted Regression

Although the global multiple regression model has the *R*-Squared value of 0.645 indicating a reasonable explanatory performance, it still leaves 35.5% of the variance in the relative risks of residential burglary unexplained. Some of this unexplained variance may be associated with the assumption of a spatial stationary process (that is, the parameters of the global model are the same for all census tracts within the study area). However, if the relationships between the socio-economic characteristics and the relative risks of residential burglary are spatially non-stationary, then the global multiple regression model will be a misspecification of the actual relationships (Fotheringham et al. 2002). One way of investigating errors in the global model is to use the Morans's *I* statistic to verify if the residuals from the global model are randomly distributed (Rogerson 2001). For the residuals from the global regression model, the value of 0.114 for Morans's *I* statistic is statistically significant at the 0.05 level and it suggests a problem of residual spatial autocorrelation. Thus, one can expect that there is an element of local variation (spatial non-stationarity) in the relationship between the relative risk for residential burglary and the neighbourhood's socio-economic characteristics.

Table 6: The ANOVA table for GWR

Source	*SS*	*df*	*MS*	*F*
OLS Residuals	55876.1	7		
GWR Improvement	12885.3	8.22	1567.271	
GWR Residuals	142990.7	54.78	784.811	1.997

Note: *SS* = sum of squares; *df* = degree of freedom, *MS* = mean sum of squares; *F* = the *F* statistic; Multiple *r*-Squared: 0.84 - 0.89.

Geographically Weighted Regression (GWR) is a method for analyzing spatial non-stationary (Brunsdon et al. 1996; Fotheringham et al. 2002). It has been developed to extend the traditional (global) regression framework by allowing local rather than global parameters to be estimated. GWR accounts for the spatial "drift" in linear relationships, by localizing the regression (placing it inside a kernel) and accepting that linearity does exist, but only over limited spatial scales. For each location in a study area, one can estimate a regression model where geographical weights are attached to observations surrounding the location. To calibrate the model, a modified weighted least squares approach is taken so that the data are weighted according to their proximity to a given point *i*. Thus the weighting of any point is not constant but varies with *i*. Data from observations closer to *i* are weighted more heavily than those from farther away. It should be noted that as well as producing localized parameter estimates, the GWR technique produces localized versions of all standard regression diagnostics (for a detailed discussion of GWR see Fotheringham et al. 2002). In this study, we use the GWR 2.1 software (Fotheringham et al. 2002) and ArcView GIS 3.2 (ESRI 1996) for calibrating the GWR model and visualize the results. The parameter estimation at any of the seventy census tracts' centroid point (or regression point) depends not only on the input data but also on the kernel chosen and the bandwidth of that kernel. In this study, a Gaussian function is selected as the kernel (that is, at the regression point the weight of a data point is unity, and it decreases with increasing distance from the regression point). A cross-validation function is used to select the bandwidth of the kernel (approximately 21 km). A

Table 7: The results of the Monte Carlo significance test for spatial variability of the GWR parameters

Parameter	*p*-value
(Intercept)	0.50
X_2	0.05
X_4	0.14
X_9	0.75
X_{10}	0.52
X_{11}	0.04
X_{12}	0.60

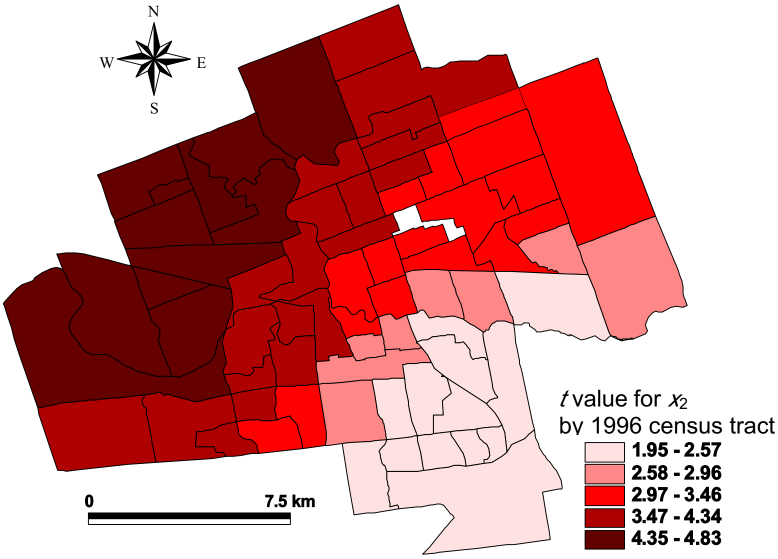

Figure 3: Spatial pattern of the local parameter estimates associated with the x_2 variable in the GWR model for London, Ontario.

Monte Carlo method is applied to carry out tests for two hypotheses: (i) that the data may be described by a global regression model rather than a global one, and (ii) whether individual regression coefficients are stable over geographic space (Fotheringham *et al.* 2002).

The results of the GWR (the F test) suggest that the GWR model is an improvement on the global model at the 0.10 level of significance (see Table 6). While the global multiple regression model accounts for 64.5% variability in the relative risk for residential burglary, the R-Squared values for the GWR models are in the range of 0.84-0.89 indicating a high explanatory performance of the models. A Monte Carlo test on the local estimates indicates that there is significant spatial variation in the local parameter estimates for the x_2 and x_{11} variables (the percentage of population with less than a grade nine education and the percentage of population without income, respectively). The spatial variation in the remaining variables is statistically insignificant (that is, there is a reasonably high probability that the variation occurred by chance) (see Table 7).

The t-values associated with the x_2 and x_{11} variables for the global model are 4.689 and 1.951, respectively (Table 5). These values represent "average" relationships across the study area. The spatial pattern of these parameters obtained by GWR is shown in Figures 4 and 5. There are clear spatial trends in the distribution of parameters associated with the x_2 and x_{11} variables. The spatial patterns of the parameters support the hypothesis of the sectoral model of the city (Hoyt 1939; Clark 1982). Specifically, they suggest that the relationships between the pattern of residential burglary and the underlying socio-economic processes (represented by the two variables) have different form at different sectors of the city. For a given proportion of population with less than a grade nine education (x_2), the relative risk of residential burglary is highest for census tracts in north-west and west sectors (affluent areas) of the city and it declines south-eastwards and eastwards (see Figure 3). To this end, it is important to note the percentages of population with less than a grade nine education are much lower in the north-west and west sectors then in the remaining parts of the city. Thus, this study suggests that an increase in the proportion of uneducated population would result in considerably higher increase in the relative risks of residential burglary in the affluent areas than in the economically disadvantaged parts of the city.

The spatial distribution of the population without income parameter shows a south-north trend (Figure 4). The proportion of the population without income has the greatest effect on the relative risk of residential burglary in the south sections of the city and the effect declines northwards. This socio-economic characteristic has the smallest effect on the relative risk of residential burglary in the north-west sector of the city. This finding suggests that an increase in the proportion of the population without income would have a considerably greater impact on the increase of relative risk ratios in the more economically disadvantaged areas of the south section of the city than in the affluent areas of the north end, especially, the north-west sector of London, Ontario.

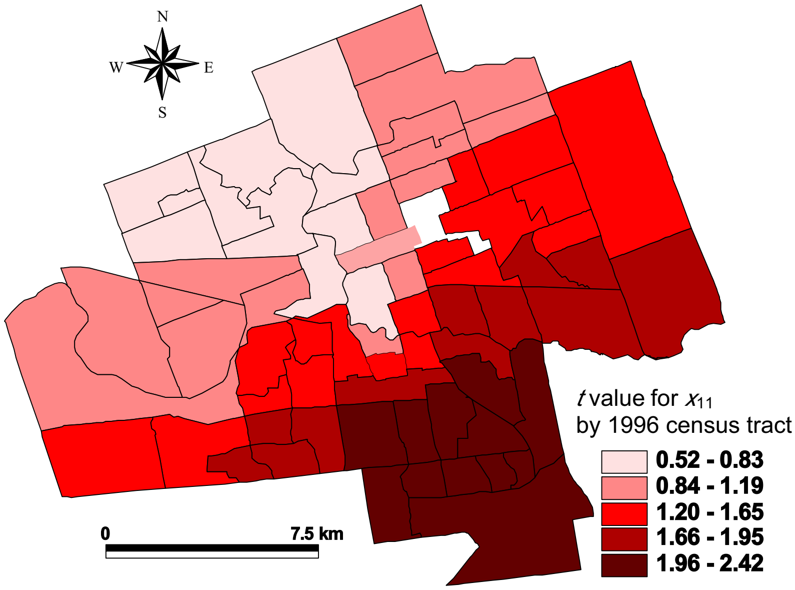

Figure 4: Spatial pattern of the local parameter estimates associated with the x_{11} variable in the GWR model for London, Ontario.

Conclusions

This paper has focused on examining the spatial pattern of relative risks of residential burglary and on analysing the relationships between the pattern and socio-economic characteristics of neighbourhoods in London, Ontario. The spatial pattern of relative risk resembles the zonal/sectoral model. The core area of the city shows the highest concentration of residential burglaries and the relative risk of burglaries tend to decline along with increasing distance from the city's center. Another distinctive feature of the spatial pattern of residential burglaries in London is the west-east division. The east section of the city is characterized by higher risks of residential burglaries.

This study has demonstrated that there are significant statistical relationships between the spatial pattern of relative risks for residential burglary and socio-economic characteristics of neighbourhoods in London, Ontario. These relationships have been examined using the standard multiple regression and geographically weighted regression (GWR). Although the global multiple regression model has a reasonable explanatory performance, the GWR analysis improves the performance of the global model and it provides an insight into the nature of spatial variations in the relationships between the relative risks of residential burglaries and socio-economic characteristics. Specifically, the study has demonstrated that there are significant spatial variations of the relationships between the relative risks of residential burglaries and such socio-economic variables as the percentage of population with less than a grade nine education and the percentage of population without income. The results of this research suggest that preventive policies should be informed by an understanding the crime contextual factors. In particular, these factors must be examined locally and *different policies* aimed at preventing and reducing crime should be applied to *different sectors* of the city. This observation is based on the offender–residence hypothesis: burglars tend to commit an offence in their neighbourhoods (Herbert and Hyde 1985; Ceccato et al. 2002, Wright and Decker 1996). However, this issue requires further examination that involves using data on the actual location of offenders in the study area.

It is important to emphasize that some of the findings of this research have supported the literature, while others have contradicted it. There are many controversies and conflicting results in previous studies concerning crime patterns and the relationships between the level of crime and contextual variables (Davidson 1981; Bursik and Grasmick 1993). In particular, more research is needed to examine the relationship between income, unemployment, and education variables and the level of crime. This is a complex issue. Some studies suggest that low educational achievement causes high unemployment and low income, which in turn motivates crime. This study has revealed that there is a significant relationship between crime and educational level, while the level of unemployment is insignificantly related to the relative risks of residential burglaries. One can argue that these inconsistent findings might be due to the way in which neighbourhoods are defined. Patterns of crime may vary widely depending on whether the areas studied are locally acknowledged neighbourhoods, official administrative areas, or electoral districts (Bursik and Grasmick 1993; LaVigne and Wartell 1998; Harries 1999).

Acknowledgements

The authors would like to thank the London Police Department for generously providing their data and Annette Swalwell (Senior Business Systems Analyst in the Department) for helpful comments on the data. Also, we wish to thank two anonymous reviewers for their insightful comments and suggestions. All errors remain those of the authors.

References

Anselin, L., J. Cohen, D.Cook, W. Gorr, and G. Tita. 2000. Spatial analyses of crime *Criminal Justice* 4: 213-262.

Bailey, T.C. and A.C. Gatrell. 1995. *Interactive Spatial Data Analysis*. New York, John Wiley & Sons Inc.

Block, C. R., M. Dabdoub, and S. Fregly. 1995. *Crime Analysis Through Computer Mapping*. Washington, D.C. Police Executive Research Forum.

Bowers, K. and A. Hirschfield. 1999. Exploring links between crime and disadvantage in North-West England: An analysis using Geographical Information Systems. *International Journal of Geographical Information Science* 13: 159-184.

Brantingham, P.J. and P.L. Brantingham. 1991. *Environmental Criminology*. Prospect Heights, IL, Waveland

Brunsdon, C., A. S. Fotheringham, and M. E. Charlton. 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis* 28: 281-298.

Burgess, E.W. 1929. Urban Areas. In T.V. Smith and L.D. White (Eds.) *Chicago: An Experiment in Social Science Research* Chicago, IL: University of Chicago Press. pp. 113-138).

Bursik, R. J. 1993. *Neighborhoods and Crime: The Dimensions of Effective Community Control*. New York, Macmillan.

Bursik, R. J. 1988. Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects. *Criminology* 26:519-51.

Bursik, R. J., Jr., and H. G. Grasmick, 1993. Economic deprivation and neighborhood crime rates, 1960-1980. Law and Society Review*, 27:* 263-268.

Ceccato, V., R. Haining, and P. Signoretta. 2002. Exploring Offence Statistics In Stockholm City Using Spatial Analysis Tools. *Annals of the Association of American Geographers* 92: 29-51.

Clark, D. 1982. *Urban Geography.* London, Billing and Sons Limited.

Craglia, M., R. Haining and P. Wiles. 2000. Comparative Evaluation of Approaches to Crime Pattern Analysis *Urban Studies* 37(4): 711–729.

Davidson, R.N. 1981. *Crime and Environment*, London, Croom Helm.

ESRI (1996) *Using ArcView GIS*, Redlands, CA., Environmental Systems Research Institute Inc.

Fotheringham, A. S., M. E. Charlton, and C. Brunsdon. 2002. *Geographically Weighted Regression the analysis of spatially varying relationships*. London, John Wiley & Sons, Ltd.

Haining R. 2003. *Spatial Data Analysis: Theory and Practice*, Cambridge, UK: Cambridge University Press.

Harries, K. 1999. *Mapping Crime: Principle and Practice*, Washington DC, Crime Mapping Research Center, National Institute of Justice.

Herbert, D. T., and S.W. Hyde. 1985. Environmental criminology: Testing some area hypotheses. *Transactions of the Institute of British Geographers* **10**: 259–274.

Hirschfield, A., Brown P., and P. Todd (1995) GIS and the analysis of spatially-referenced crime data: Experiences in Merseyside, U.K. *International Journal of Geographical Information Systems.* 9, 191-210.

Hoyt, H. (1939). *The Structure and Growth of Residential Neighbourhoods in American Cities.* Washington D.C.: Federal Housing Administration.

Insightful Corporation. 2001. *S-Plus 6 for Windows: User's Guide*. Seattle, WA, Insightful Corporation.

Jenks, G. F. and F. C. Caspall. 1971. Error on choroplethic maps: Definition, measurement, reduction. *Annals of the Association of American Geographers* 61(2): 217-244.

Kohfeld, C.W. and J. Sprague. 1988. Urban Unemployment Drives Crime. *Urban Affairs Quarterly.* 24: 215-241.

LaVigne, N. and J. Wartell (1998) *Crime Mapping Case Studies: Successes in the Field*, Police Executive Research Forum, Washington, D.C.

Marcuse, P. 1993. What's So New About Divided Cities? *International Journal of Urban and Regional Research* 17: 355-65.

Messner S. F., L. Anselin, D. F. Hawkins, G. Deane, S. E. Tolnay, and R. D. Baller. 1998. An Atlas of the Spatial Patterning of County-Level Homicide, 1960-1990, The 50th Annual Meeting of the American Society of Criminology, November 11-14, 1998, Washington, D.C.

Miller, A. J. 1990. *Subset Selection in Regression*. New York, Chapman and Hall.

Ormsby, T. and J. Alvi. 1999. *Extending ArcView GIS: with Network Analyst, Spatial Analyst and 3D Analyst*, Redlands, ESRI Press.

Paternoster R., and S.D Bushway. (2001) Theoretical and empirical work on the relationship between unemployment and crime. *Journal of Quantitative Criminology.* 17: 391-407.

Poetz, A. 2003. Spatial Patterns of Residential Burglaries in London, Ontario. Unpublished B.A. Thesis, University of Western Ontario, London, Ontario.

Ratcliffe, J.H., and M. J. McCullagh. 1999. Hotbeds of crime and the search for spatial accuracy. *Journal of Geographic Systems* 1: 385-398.

Reilly B., and Witt, R. (1992) Crime and unemployment in Scotland: An econometric analysis using regional data. *Scottish Journal of Political Economy.* 39: 213-228.

Rogerson, P. A. (2001) *Statistical Methods for Geography*, London, SAGE Publications.

Sampson, R. J., and W. B. Groves.1989. Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology* 94: 774–802.

Selvin, S. 1998. *Modern Applied Biostatistical Methods Using S-Plus*.New York, Oxford University Press.

Shaw, C.R. and H. D. McKay. 1972. *Juvenile Delinquency and Urban Areas*. Chicago, University of Chicago Press.

Smith, D. A. and G. R. Jarjoura. 1989. Household characteristics, neighborhood composition and victimization risk. *Social Forces* 68: 621–640.

Statistics Canada. 1996. Canada – Census – 1996. Data located at Internet Data Library System, The University of Western Ontario, http://www.ssc.uwo.ca/idlsv2/

Swalwell, A. 2002. Personal Communication. Senior Business Systems Analyst, London (Ontario) Police.

Timbrell, M. (1990) Does unemployment lead to crime? *Journal of Interdisciplinary Economics*. 3: 223-242.

Van Kempen, E.T. 1994 . The Dual City and The Poor: Social Polarisation, Social Segregation and Life Chances. *Urban Studies* 31: 995-1015.

Weisburd D. and T. McEwen. 1998. *Crime Mapping and Crime Prevention,* Monsey, New York: Criminal Justice Press.

Wiles, P. and A. Costello. (2000). *The road to nowhere*: *The evidence for traveling criminals*. Home Office Research Study. London: Home Office.

Wright R. and S. Decker. 1994. *Burglars on the Job.* Boston: Northeastern University Press.