SpAM

SpAM (Spatial Analysis and Methods) presents short articles on the use of spatial statistical techniques for housing or urban development research. Through this department of Cityscape, *the Office of Policy Development and Research introduces readers to the use of emerging spatial data analysis methods or techniques for measuring geographic relationships in research data. Researchers increasingly use these new techniques to enhance their understanding of urban patterns but often do not have access to short demonstration articles for applied guidance. If you have an idea for an article of no more than 3,000 words presenting an applied spatial data analysis method or technique, please send a one-paragraph abstract to rwilson@umbc.edu for review.*

Predicting Local Crime Clusters Using (Multinomial) Logistic Regression

Martin A. Andresen
Simon Fraser University

Abstract

Understanding hot spots of crime has been a concern of spatial criminology for nearly 200 years. A number of methods are used to identify and calculate hot spots, such as dot maps, surface interpolation (kernel density estimation), and statistically identified cluster analysis. Relating to the latter set of methods, local Moran's I *is one of the more commonly used methods for identifying local crime clusters. One important aspect for this method for subsequent analysis is that it uses areas, such as census boundary units, to identify local clusters of crime. Consequently, census data may be used to predict and better understand these local crime clusters. In this article, I use multinomial logistic regression and census variables to predict the local crime clusters identified by local Moran's* I. *This analysis shows a number of nuances regarding local crime clusters, and the spatial patterns of crime, more generally, can be identified using this two-stage approach. Such an approach provides a better understanding of spatial crime patterns than the more common regression methods.*

Introduction

The use of spatial statistical methods in criminology is becoming increasingly popular with the recognition that spatial relationships exist in crime data that complicate traditional statistical estimation. In addition, spatial statistical techniques are becoming applied more because of their ability to ask new questions to better understand the criminal event. One of these statistical techniques is explicitly spatial and considers local relationships: local crime clusters using local Moran's *I*. This technique and others like it are "local" because, rather than calculating a statistic for the entire study area (a "global" statistic), it calculates a statistic for each unit under analysis (census tracts, for example).

Although instructive on their own, simply with their identification, a better understanding of the determinants of these local crime clusters is of interest. This article discusses a two-stage research methodology that incorporates local crime clusters and their prediction. The two-stage approach first identifies the local crime clusters in the study area and then a multinomial logistic regression to identify the predictors for each local crime cluster type. This two-stage approach shows that the output from a local Moran's *I* and multinomial logistic regression is able to identify factors specific to each local crime cluster type beyond what would be found in more common regression analyses.

Data and Methods

The data for the following analyses come from Vancouver, British Columbia, Canada. Vancouver, a city on the west coast of Canada, just north of the U.S. border with Washington State, is Canada's third largest census metropolitan area (CMA), with a population that is slightly more than 2.3 million. Although Vancouver is the third largest CMA in Canada, it historically has had the highest crime rate for CMAs in the country. In 2013, for example, Vancouver's total crime rate (excluding traffic) was 6,897 per 100,000 residents, more than double the crime rate in the largest Canadian CMA, Toronto, at 2,941 per 100,000 (Boyce, Cotter, and Perreault, 2014).

Data and Spatial Units of Analysis

The spatial unit of analysis for Vancouver is the dissemination area (DA). The DA, approximately the size of a census block group in the U.S. census, contains approximately 400–700 people and is composed of one or more blocks. In Vancouver, 990 DAs are used for our analyses.

Crime data for Vancouver consist of calls for service made to the Vancouver Police Department for the year 2001. The calls-for-service database is the set of calls made to the Vancouver Police Department directly, calls allocated to them through the 911 emergency system, and calls made by police while on duty. These data consist of automotive theft (theft of vehicle and theft from vehicle), burglary (commercial and residential), and violent crime (assault, fighting, holdups, homicide, robbery, sexual assault, and stabbing). Each call for service includes the complaint code/description, listed previously, and the location in the form of an address of the call for subsequent mapping and spatial analysis. These data were geocoded to the street network with a 94-percent success rate, exceeding the minimum acceptable hit rate of 85 percent set by Ratcliffe (2004).

The census data used in the inferential analyses represent the appropriate years of crime data for Vancouver (2001) at the dissemination area level. These explanatory variables were chosen, in both cases, to represent social disorganization theory and routine activity theory (Cohen and Felson, 1979; Shaw and McKay, 1942). For the Vancouver analyses, the following variables were employed: population change, percent; males ages 15 to 24, percent; single-parent families, percent; recent immigrants, percent; ethnic diversity, measured using the Blau Index; unemployment, percent; post-secondary completion, percent; average income in thousands of dollars; population density; average dwelling value in thousands of dollars; rentals, percent; and housing in major repair, percent.

Local Indicator of Spatial Association, Local Moran's *I*

As mentioned previously, local Moran's *I* has been used in a number of criminological contexts to represent local clusters of crime. Local Moran's *I* is classified as a local indicator of spatial association (LISA) because it is mathematically related to the global spatial statistic, Moran's *I* (Anselin, 1995). This local spatial statistic identifies spatial clustering at the local level, and it does so for each spatial unit of analysis, indicating if each spatial unit of analysis is surrounded by similar or dissimilar values.

The local Moran's *I* statistic is calculated in the following manner—

$$I_i = \frac{\left(x_i - x^*\right)\sum_j w_{ij}\left(x_j - x^*\right)}{\sum_i \left(x_i - x^*\right)^2 / n}, \tag{1}$$

where x_i is the value of variable x in spatial unit i, x^* is the mean of x, n is the number of spatial units, and w_{ij} is the spatial weights matrix that measures the strength of the relationship between two spatial units. In the analyses that follow, spatial weights are defined using first order Queen's contiguity, such that all spatial units that are contiguous are considered neighbors, even if they touch at only a corner. The local Moran's *I* statistic ranges from -1 (perfect negative spatial autocorrelation) to +1 (perfect positive spatial autocorrelation). These values are then used for the following classifications of local clusters: high-high, low-low, low-high, and high-low. High-high and low-low represent local positive spatial autocorrelation, high crime-rate areas surrounded by other high crime-rate areas (hot spots of crime) and low crime-rate areas surrounded by other low crime-rate areas (cool spots of crime), respectively. Low-high and high-low represent local negative spatial autocorrelation, low crime-rate areas surrounded by high crime-rate areas and high crime-rate areas surrounded by low crime-rate areas, respectively. A final and fifth classification also represents no statistically significant spatial clustering.

Maps of these local crime clusters are shown in the Presence and Prediction of Local Crime Clusters section, but these local crime cluster classifications can also be used as the dependent variable in a regression context. As such, we can use a regression technique—(multinomial) logistic regression—to predict the respective categories.

(Multinomial) Logistic Regression

Because of the categorical nature of the local crime cluster classifications, a logistic regression model is appropriate (McFadden, 1974, 1981). It is important to note, however, that just because it is possible to have a total of five local crime cluster classifications, it does not mean that all will be present in any given analysis. In the research on Vancouver, as shown in the following section, all five classifications are present.

Multinomial logistic regression is a statistical technique that specifies the dependent variable as a category, rather than as a continuous or count-based variable. The local crime clusters that emerge from local Moran's *I* is an example of such data. The output from the multinomial logistic regression is a set of parameters for each local crime cluster (no statistically significant clustering is used as the base category in the regression model) that can then be used to calculate the probability that each local crime cluster will occur, given the values of the estimated parameters and the explanatory variables.

The formula for the multinomial logistic model is as follows—

$$\text{Prob}\,(Y = J) = \frac{e^{X\beta_i}}{1 + \sum_{i=1}^{J} e^{X\beta_i}}, \tag{2}$$

where Y is the local cluster type, $J + 1$ is the number of alternatives, e is the natural exponential function, X is the matrix of independent variables, and β_i are estimated parameters. In the analyses that follow, all variables listed previously are included in the specification of equation 2. All statistically insignificant variables subsequently were removed using *t*-tests for individual variables and likelihood-ratio tests for joint significance tests. This step is done to avoid removing relevant variables because of multicollinearity.

One final consideration for the multinomial logistic regression model is how to assess the individual impact of explanatory variables on the probability that each local crime cluster will occur. This assessment can be done by calculating marginal effects or using odds-ratios. Because of the nonlinearity of the multinomial logistic functional form, the marginal effect of the explanatory variable on the outcome variable is not constant. Rather, the marginal effect is a function of all the estimated coefficients in the model (Greene, 2000; Kennedy, 2003; Wooldridge, 2002). For example, the change in probability for choice J is the partial derivative of the likelihood function with respect to x_i—

$$\text{Marginal Effect} = \beta_i e^{X\beta}\left(1 + e^{X\beta}\right)^{-2}. \tag{3}$$

Clearly evident from equation 3 is that the partial derivative varies with the values of *all* the explanatory variables in the estimated model. As such, a choice needs to be made regarding which values for the explanatory variables should be used to calculate the marginal effects. The most common method of calculating the marginal effects is to use the average values for each explanatory variable (Kennedy, 2003). The marginal effects shown in the tables that follow are calculated in this manner.

The odds-ratio is the exponential function of the estimated parameter, e^{β}. This odds-ratio represents the relative change in the probability of a local crime cluster when one unit increases the value of the explanatory variable. For example, if the odds-ratio is 1.20, then a one-unit increase in an explanatory variable leads to a 20-percent increase in the probability of a local crime cluster; similarly, if the odds-ratio is 0.80, then a one-unit increase in an explanatory variable leads to a 20-percent decrease in the probability of a local crime cluster. An example of each method is provided in the following section for the Vancouver analyses. All estimation was undertaken using R: A Language and Environment for Statistical Computing (http://www.R-project.org).

Presence and Prediction of Local Crime Clusters

The results for the local crime clusters in Vancouver are shown in exhibits 1 through 4. Exhibit 1 shows the counts and percentages of local crime cluster types for the dissemination areas. Immediately evident is the insignificant clustering that is, by far, the most common result for Vancouver dissemination areas, 77.0 to 83.0 percent. This exhibit clearly shows that hot and cool spots of crime are far from a dichotomy. The high-high clusters account for 4.0 to 8.5 percent of the dissemination areas, whereas low-low clusters account for 8.0 to 14.0 percent of the dissemination areas. As such, in Vancouver, statistically defined cool spots of crime are uncommon but more present than hot spots of crime. The remaining two local crime cluster types, low-high and high-low, represent negative spatial autocorrelation and are rather rare, 1.0 to 2.5 percent of dissemination areas.

Turning to the maps, exhibit 2 shows the local crime clusters for automotive theft in Vancouver. The high-high crime cluster is primarily in the downtown peninsula of Vancouver. The areas to the east and immediately south of the downtown peninsula are Downtown Eastside and False Creek, respectively. These areas of Vancouver have high crime levels, particularly Downtown Eastside, and have not surprisingly been identified as such. The low-low local crime clusters are primarily on the west side of the city, historically the wealthiest area of Vancouver. As with the hot spots of crime, this effect is not a surprise. The negative spatial autocorrelation local crime clusters, high-low and low-high, are scattered around the city. Perhaps most interesting is the presence of low-high local crime clusters in the Downtown Eastside and False Creek areas. These clusters represent "havens" from automotive theft in a crime hot spot.

Exhibit 1

Percentages and Counts of Cluster Types, Dissemination Areas, Vancouver, British Columbia, Canada, 2001

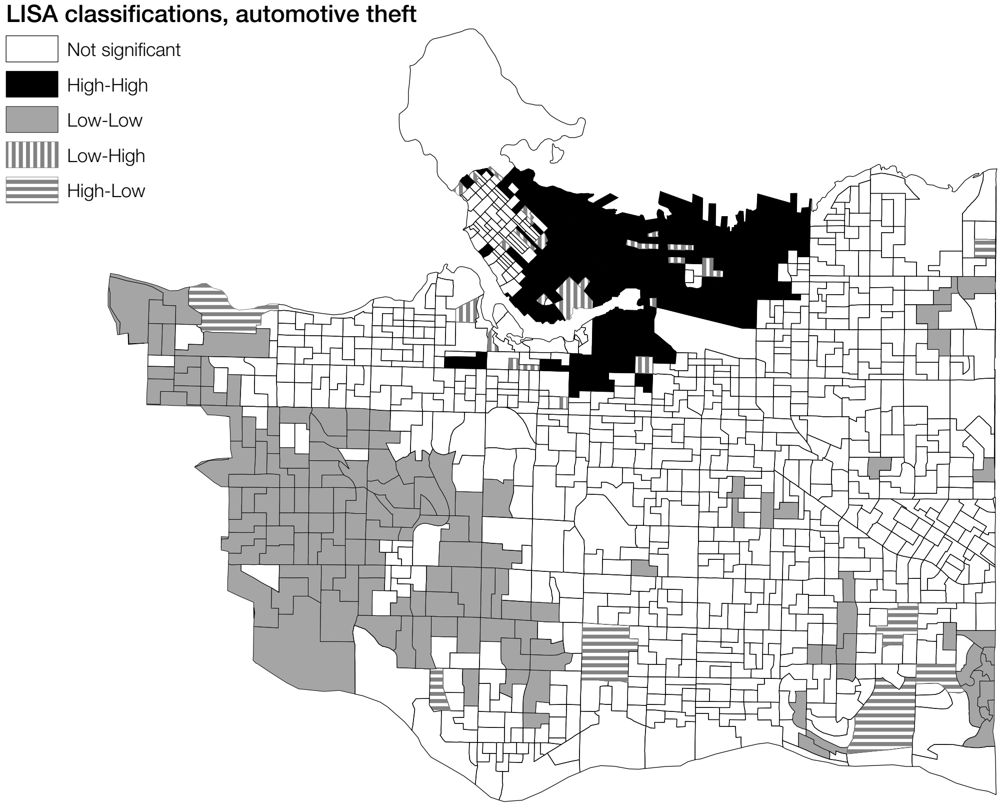
	Cluster Type				
	Insignificant	**High-High**	**Low-Low**	**Low-High**	**High-Low**
Automotive theft	77.2	5.6	14.0	2.5	0.7
	(764)	(55)	(139)	(25)	(7)
Burglary	76.7	8.6	8.3	3.9	2.5
	(759)	(85)	(82)	(39)	(25)
Violent crime	82.9	4.0	10.7	1.6	0.7
	(821)	(40)	(106)	(16)	(7)

Note: Counts are reported in parentheses.

Source: Andresen (2011)

Exhibit 2

LISA Map, Dissemination Areas, Automotive Theft, Vancouver, British Columbia, Canada

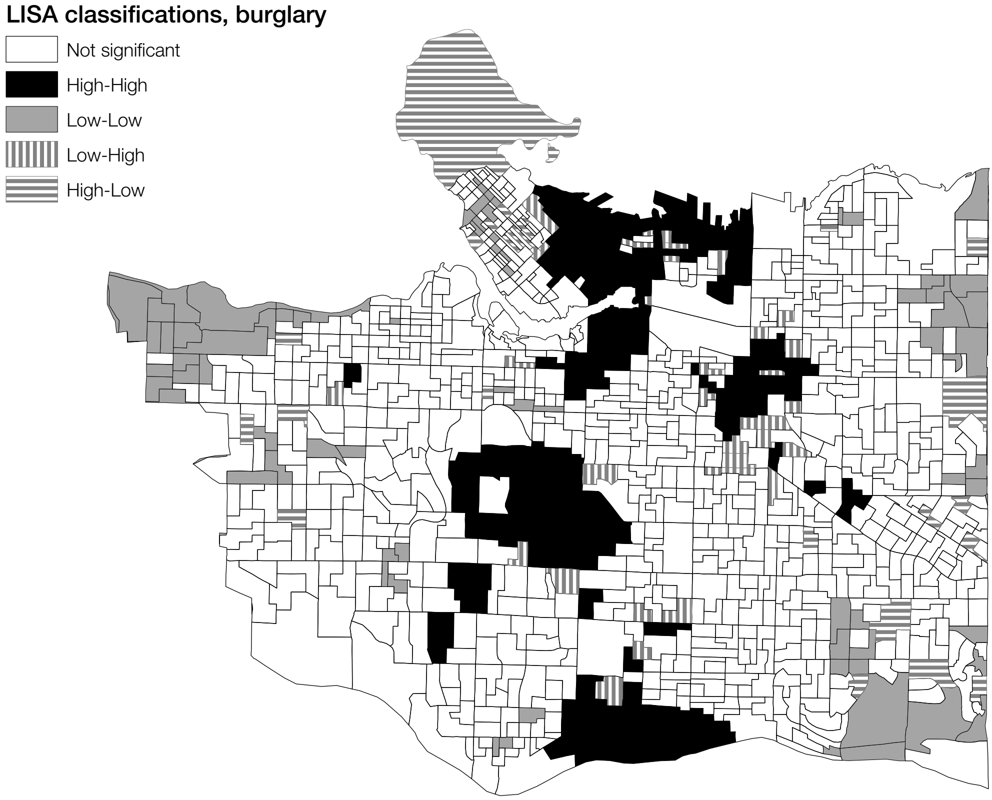


LISA = local indicator of spatial association.
Source: Andresen (2011)

The local crime clusters for burglary in Vancouver, exhibit 3, are in quite different areas of the city, showing the importance of disaggregating crime types, particularly for spatial analyses (Andresen and Linning, 2012). Hot spots of burglary are still present in the poorer areas within the downtown peninsula and the Downtown Eastside. This result is hardly surprising. A hot spot of burglary is in the center of Vancouver, however, in an area that would not be expected. Although the far eastern portion of this hot spot is close to an area that traditionally has higher crime in Vancouver, much of the rest of this hot spot is not. The remaining hot spots are primarily on the east side of Vancouver, traditionally a lower income area of the city. Within the east side of the city, however, some areas have relative affluence; one of these areas has been shown to be a low-low local crime cluster in the southeast corner of the city.

Exhibit 3

LISA Map, Dissemination Areas, Burglary, Vancouver, British Columbia, Canada



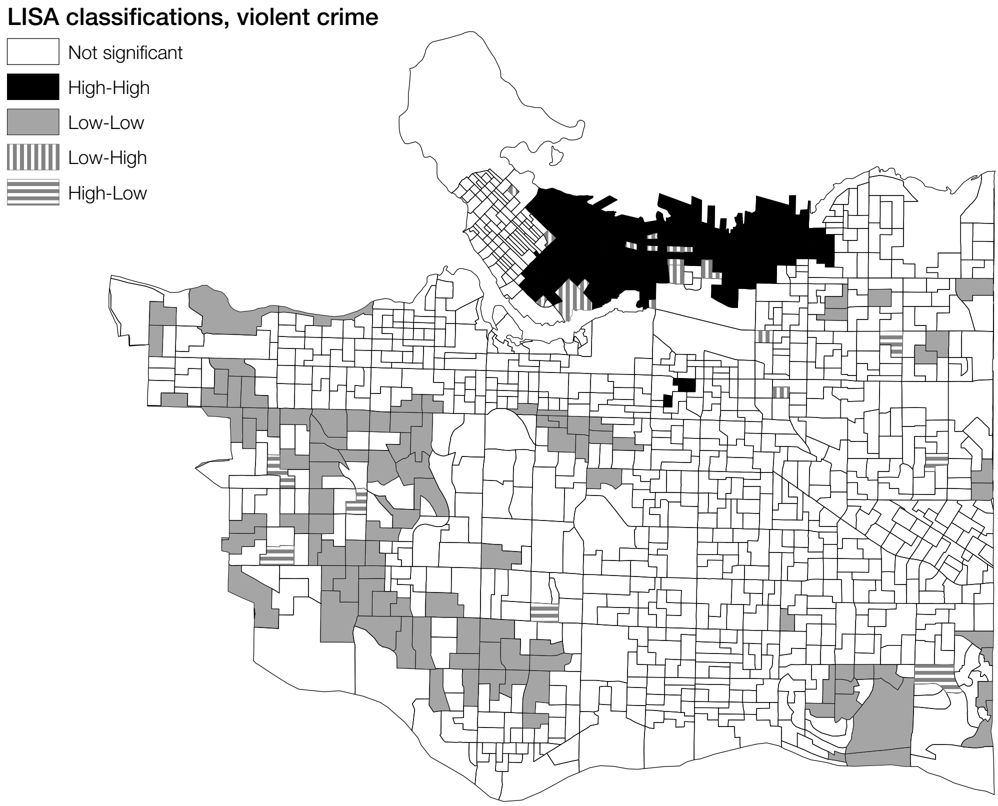
LISA = local indicator of spatial association.
Source: Andresen (2011)

Exhibit 4 shows the local crime clusters for violent crime in Vancouver. The overall pattern of automotive theft: low-low local crime cluster in the western portion of the city with a high-high local crime cluster in the downtown peninsula and the Downtown Eastside. One notable difference regarding the violent crime high-high local crime cluster is that it does not extend south like automotive theft but extends farther east along a main thoroughfare in the city—Hastings Street.

The local crime clusters are interesting on their own, but predicting their presence is of greater interest here. Exhibit 5 shows the estimation results from the multinomial logistic regression for automotive theft. We will not go into any detail regarding the magnitudes of the parameters and theoretical predictions here—the reader is referred to Andresen (2011) for such discussions. The most interesting comparisons for our purposes here are high-high versus low-low and low-high versus high-low. What is important to note is that each of these comparisons is not as expected. The most obvious is high-high and low-low, representing hot spots and cool spots of crime, respectively. One normally would expect the same variables to matter for each area but to have the opposite sign. This expectation does not occur here, although some overlap exists. Five explanatory variables

Exhibit 4

LISA Map, Dissemination Areas, Violent Crime, Vancouver, British Columbia, Canada



LISA = local indicator of spatial association.
Source: Andresen (2011)

do not have a corresponding presence for the other local crime cluster type, and in two cases the signs on the estimated parameters are the same. This result is more pronounced for a comparison between low-high and high-low, having only one common estimated parameter.

Perhaps the most instructive aspects in exhibit 5, and in the subsequent tables as well, are the differences between the calculated marginal effects (middle number inside the brackets) and the odds-ratios (number to the right of the brackets), with the latter being more commonly reported in the literature. According to the odds-ratios, some of the explanatory variables have large impacts on the probability of local crime clusters. For example, in the case of high-high, young males, single parents, and the unemployment rate all appear to have a large impact on the probability: 16- to 18-percent change in probabilities. Remember, however, that these are relative changes that depend on the baseline probability of an event such as a local crime cluster occurring. The marginal effects show that the actual impacts on the probability that a high-high local crime cluster will occur are actually rather small, one-tenth of 1 percent. Such big differences in relative and absolute changes in probability are present for all other local crime cluster types as well. This analysis shows the importance of calculating marginal effects in a (multinomial) logistic regression context.

Exhibit 5

Multinomial Logistic Regression Results, Dissemination Areas, Automotive Theft, Vancouver, British Columbia, Canada

	High-High	Low-Low	Low-High	High-Low
Constant	– 3.848*	– 3.954*	– 6.191*	
Population change (%)		– 0.040 [– 0.240]* 0.96	0.052 [0.051] 1.05	– 0.189 [– 0.000]* 0.83
Males ages 15–24 (%)	– 0.180 [– 0.119] 0.84	0.252 [1.533]* 1.29		
Single-parent families (%)	– 0.198 [– 0.130]* 0.82	0.098 [0.597]* 1.10		
Ethnic diversity	0.029 [0.019] 1.03	– 0.028 [– 0.171]* 0.97		
Unemployment rate	0.159 [0.105]* 1.17		0.116 [0.114]* 1.12	
Post-secondary completion (%)	0.031 [0.021]* 1.03			
Average income ($ thousands)[a]	0.027 [0.018]* 1.03	0.017 [0.105]* 1.02	0.022 [0.021]* 1.02	
Population density	– 0.006 [– 0.004]* 0.99	0.001 [0.001]* 1.00		– 0.127 [– 0.000]* 0.88
Dwelling value ($ thousands)[a]	– 0.011 [– 0.007]* 0.99		– 0.005 [– 0.005] 0.99	
Rentals (%)		– 0.018 [– 0.106]* 0.98		
Housing in major repair (%)	– 0.091 [– 0.060]* 0.91	– 0.057 [– 0.349]* 0.95		
Probability of cluster	0.66	6.08	0.99	0.00
Pseudo-R^2	0.337			
Percent correct	82.32			

* $p < 0.05$.
[a] *In 2001 Canadian dollars.*
Notes: Marginal effects, shown in brackets, are calculated using average values. Odds-ratios are below the marginal effect. All retained variables are statistically significant at least at the 10-percent level.

We can also see that the estimated probability, using the average values of the explanatory variables, of any of these local crime clusters is rather low. Only the cluster type low-low has a probability greater than 1 percent of occurring. Finally, the goodness-of-fit for the multinomial logistic regression model is moderate with a pseudo-R^2 of 0.337 and 82.32 percent of its estimates correctly identifying the appropriate local crime cluster type.

Turning to the results for burglary, exhibit 6, the same basic results are present, but with fewer variables: different local crime cluster types have different sets of statistically significant explanatory variables, aside from one case—the odds-ratios do not indicate large magnitude impacts on the prediction of local crime clusters. Notably different is that the probability that each local crime cluster will occur is greater than for automotive theft, except for low-low. In fact, both high-high and low-low have probabilities of occurring that are approximately 6 percent. The goodness-of-fit measures, however, are not as strong as for automotive theft: a pseudo-R^2 of 0.074 and 76.67 percent of its estimates correctly identifying the appropriate local crime cluster type.

Exhibit 6

Multinomial Logistic Regression Results, Dissemination Areas, Burglary, Vancouver, British Columbia, Canada

	High-High	Low-Low	Low-High	High-Low
Constant	– 5.26*		– 2.386*	
Population change (%)	0.026 [0.158]* 1.03			
Males ages 15–24 (%)				0.291 [0.379]* 1.34
Recent immigrants (%)		– 0.031 [– 0.180]* 0.97		– 0.060 [– 0.079]* 0.94
Ethnic diversity	0.025 [0.152]* 1.03	– 0.019 [– 0.110] 0.98		– 0.031 [– 0.040] 0.97
Unemployment rate	0.072 [0.435]* 1.08		0.067 [0.153]* 0.94	
Dwelling value ($ thousands)[a]	0.003 [0.016]* 1.00	0.001 [0.007] 1.00	– 0.004 [– 0.009] 0.99	– 0.007 [– 0.009]* 0.99
Housing in major repair (%)		– 0.042 [– 0.244]* 0.96		
Probability of cluster	6.06	5.83	2.31	1.30
Pseudo-R^2	0.074			
Percent correct	76.67			

* $p < 0.05$.
[a] *In 2001 Canadian dollars.*
Notes: Marginal effects, shown in brackets, are calculated using average values. Odds-ratios are below the marginal effect. All retained variables are statistically significant at least at the 10-percent level.

The results for violent crime are more similar to those for automotive theft. More statistically significant explanatory variables are retained for this model, particularly for high-high and low-low. Moreover, as with the other two crime types, each local crime cluster type retains a different set of explanatory variables and, when a variable is present in both high-high and low-low, the estimated parameters are opposite in sign, as would be expected. As with automotive theft, the odds-ratios for violent crime local crime clusters are, at times, indicating a large magnitude impact with the actual probability changes being small. The goodness-of-fit values are moderate with a pseudo-R^2 of 0.255 and 83.43 percent of its estimates correctly identifying the appropriate local crime cluster type.

Exhibit 7

Multinomial Logistic Regression Results, Dissemination Areas, Violent Crime, Vancouver, British Columbia, Canada

	High-High	Low-Low	Low-High	High-Low
Constant		– 1.869*	– 6.597*	
Males ages 15–24 (%)		0.116 [0.719]* 1.12		
Single-parent families (%)	– 0.195 [– 0.046]* 0.82			0.376 [0.000]* 1.46
Ethnic diversity	0.032 [0.008] 1.03	– 0.026 [– 0.158]* 0.97		
Unemployment rate	0.161 [0.038]* 1.18	– 0.048 [– 0.299]* 0.95	0.117 [0.066]* 1.12	
Average income ($ thousands)[a]	0.022 [0.005]* 1.02			
Population density	– 0.006 [– 0.001]* 0.99	0.001 [0.002]* 1.00		– 0.052 [– 0.000]* 0.95
Dwelling value ($ thousands)[a]	– 0.018 [– 0.004]* 0.98	0.002 [0.009] 1.00		
Rentals (%)		– 0.020 [– 0.125]* 0.98	0.031 [0.017]* 1.03	– 0.054 [– 0.000] 0.95
Housing in major repair (%)	– 0.078 [– 0.018]* 0.93			
Probability of cluster	0.24	6.20	0.56	0.00
Pseudo-R^2	0.255			
Percent correct	83.43			

* $p < 0.05$.
[a] *In 2001 Canadian dollars.*
Notes: Marginal effects, shown in brackets, are calculated using average values. Odds-ratios are below the marginal effect. All retained variables are statistically significant at least at the 10-percent level.

Conclusion

This article investigates not only the use of local indicators of spatial association, specifically the local crime clusters of local Moran's *I*, but also their prediction. Although such spatial statistical techniques are useful and interesting on their own, it is also important to better understand which variables are associated with each local crime cluster so that we may be able to increase their presence (low-low) or decrease their presence (high-high).

The advantage of using this two-stage approach versus an ordinary least squares (OLS) or spatial regression of crime rates is that the local crime cluster technique specifically identifies the areas of interest: hot spots of crime and cool spots of crime. Then, the multinomial logistic regression can estimate the set of parameters that matter specifically for each local crime cluster type, whereas an OLS or spatial regression will estimate one parameter to be used for all areas. In addition, this estimation technique does not provide estimates for the places that do not exhibit spatial clustering, but these places are used as the base category for estimating the probability of the statistically significant local crime clusters.

Overall, this combination of techniques has proven instructive. This article found that different explanatory variables matter for the different local crime clusters, something that would not emerge in an OLS or spatial regression context. This result allows for a higher degree of understanding with the nuances of crime in an urban environment.

Acknowledgments

The author thanks Ronald Wilson for his discussions regarding this article.

Author

Martin A. Andresen is a professor in the School of Criminology and is the Associate Director of the Institute for Canadian Urban Research Studies at Simon Fraser University.

References

Andresen, Martin A. 2011. "Estimating the Probability of Local Crime Clusters: The Impact of Immediate Spatial Neighbors," *Journal of Criminal Justice* 39 (5): 394–404.

Andresen, Martin A., and Ehsan Jozaghi. 2012. "The Point of Diminishing Returns: An Examination of Expanding Vancouver's Insite," *Urban Studies* 49 (16): 3531–3544.

Andresen, Martin A., and Shannon J. Linning. 2012. "The (In)appropriateness of Aggregating Across Crime Types," *Applied Geography* 35 (1–2): 275–282.

Anselin, Luc. 1995. "Local Indicators of Spatial Association—LISA," *Geographical Analysis* 27 (2): 93–115.

Boyce, Jillian, Adam Cotter, and Samuel Perreault. 2014. *Police-Reported Crime Statistics in Canada, 2013*. Ottawa, Ontario, Canada: Canadian Centre for Justice Statistics, Statistics Canada.

Cohen, Lawrence, and Marcus Felson. 1979. "Social Change and Crime Rate Trends: A Routine Activity Approach," *American Sociological Review* 44 (4): 588–608.

Greene, William H. 2000. *Econometric Analysis,* 4th ed. Upper Saddle River, NJ: Prentice Hall.

Kennedy, Peter. 2003. *A Guide to Econometrics,* 5th ed. Malden, MA: Blackwell Publishing.

McFadden, Daniel. 1981. "Econometric Models of Probabilistic Choice." In *Structural Analysis of Discrete Data Using Econometric Applications*, edited by Charles F. Manski and Daniel McFadden. Cambridge, MA: MIT Press: 198–272.

———. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by Paul Zarembka. New York: Academic Press: 105–142.

Ratcliffe, Jerry H. 2004. "Geocoding Crime and a First Estimate of a Minimum Acceptable Hit Rate," *International Journal of Geographical Information Science* 18 (1): 61–72.

Shaw, Clifford R., and Henry D. McKay. 1942. *Juvenile Delinquency and Urban Areas: A Study of Rates of Delinquency in Relation to Differential Characteristics of Local Communities in American Cities*. Chicago: University of Chicago Press.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.