

Cell towers and the ambient population: A spatial analysis of disaggregated property crime

Patrick Johnson

School of Criminology, Simon Fraser University
8888 University Drive, Burnaby BC V5A 1S6, Canada

Martin A. Andresen (corresponding author)

m.andresen@griffith.edu.au

School of Criminology and Criminal Justice, Gold Coast Campus, Griffith University
Parklands Dr, Southport QLD 4215, Australia

Nick Malleon

School of Geography, University of Leeds
Leeds, UK LS2 9JT

Abstract

As a crime rate denominator, the ambient population has seen very limited use in a multivariate context. The current study employs a new measure of this population, constructed using cell tower location data from OpenCellID, to compare residential and ambient population-based crime rates. The chosen study area is Vancouver, BC, but the conclusions generalise to other administrations and the OpenCellID data have global coverage so the implications are applicable elsewhere. Five disaggregated property crime types are examined at the dissemination area level. Findings demonstrate striking differences in the spatial patterns of crime rates constructed using these two different measures of the population at risk. Multivariate results from spatial error models also highlight the substantial impact that the use of a theoretically-informed crime rate denominator can have on Pseudo R² values, variable retention, and trends in significant relationships. Implications for theory testing and policy are discussed. In particular, the results suggest that policies designed around residential-based crime rates risk having no effect, or even of increasing crime.

Keywords: ambient population; OpenCellID; population at risk; property crime; spatial analysis

Introduction

As a measure of crime, rates address the main limitation of raw counts of crime by controlling for the population at risk. For example, a major transit hub that experiences very high counts of crime may be assumed to have a crime problem (Boggs, 1965). The difficulty lies in defining this population. Almost invariably, the residential population of a given spatial unit is used as the denominator when crime rates are calculated. Still, as Harries (1991) pointed out, “the uncritical application of [the residential] population as a denominator for all crime categories may yield patterns that are at best misleading and at worst bizarre” (p. 148). In other words, it should not simply be taken for granted that the residential population provides the best representation of the population at risk for every crime type.

Over the years, researchers have employed a variety of alternative denominators to study crime and test theory (see Boggs, 1965; Lottier, 1938; Skogan, 1976). One measure that has emerged in recent years is the ambient population. Definitions of the ambient population in the literature are numerous and context dependent. Here we define the ambient population as one that captures the *number of people in a given area engaged in their day-to-day activities*. Importantly, this includes people who do their day-to-day activities at or near their homes, as well as people who travel longer distances. Prior research has consistently identified important differences between residential and ambient population-based crime rates at both the descriptive (Malleon & Andresen, 2016; Stults & Hasbrouck, 2015;) and inferential levels (Andresen, 2006a, 2011). Overall, the literature suggests that the ambient population can provide a very different perspective on environmental risk and opportunity. However, other than Andresen’s (2011) study of

aggregate violent crime in Vancouver, nearly all prior inferential research has been conducted at the neighbourhood or city level (see Hanaoka, 2018; Hipp et al., 2019 for two exceptions). Additionally, only a handful of prior studies have used ambient population-based crime rates inferentially as dependent variables (Andresen, 2006a, 2011).

The current study adds to the literature on the ambient population and crime in four ways. First, this spatial analysis of property crime in Vancouver, British Columbia was conducted at the finer, dissemination area level. Larger units, such as census tracts, often hide the heterogeneity that becomes apparent at finer resolutions (Andresen & Malleson, 2013), making it more difficult to control for the observed variations in a multivariate analysis. Second, ambient population-based disaggregated crime rates are used as dependent variables in spatial regression models. Third, relatively current data from 2016 are used. Fourth, this study employs a new measure of the ambient population calculated using open source cell tower location data, OpenCellID; these data are a global collaborative community project that locates cellular towers that can then be used to estimate where people are located. The data are available globally, so the analysis could be repeated anywhere where there are a sufficient number of smart phone users to populate the OpenCellID database and suitable crime and census data. Using the frameworks and constructs of social disorganization theory (physical status, economic status, and population characteristics) and routine activity theory (motivated offenders, suitable targets, and lack of capable guardians), these spatial analyses examine whether or not there are important differences between regression models using either residential or ambient population-based crime rates as dependent variables.

Related research

Alternative denominators have been used in the study of crime at least as far back as 1938, when Lottier examined state-by-state differences for a variety of crime types. When calculating auto theft rates, Lottier (1938) used the number of automobiles registered in the state as the population at risk. While Lottier (1938) did not discuss the reasoning behind his decision to use this denominator, doing so would suggest that he felt this measure would better capture environmental risk for this crime type than the residential population. Indeed, with lower rates of car ownership during the 1930s, a rate based on the residential population would likely have been low and would not have provided an accurate indication of the risk facing automobile owners.

The first study that examined and compared crime rates with alternative denominators was Boggs (1965), who noted that crime rates based on the residential population may lead to spuriously high rates for central business districts, that often have few residents, but “large numbers of such targets as merchandise on display, untended parked cars on lots, people on the streets, money in circulation, and the like” (p. 900). Overall, Boggs (1965) found that some of the traditional, residential population-based crime rates were highly correlated with their alternative counterparts: homicide and aggravated assault, forcible rape, and residential day burglary, suggesting that alternative denominators were likely not of any particular value. Other crime types had lower correlation coefficients between the traditional and crime-specific rates, such as auto theft for joy riding and business robbery. Interestingly, three crime types had negative correlations between the standard and alternative rates: non-residential night burglary, non-residential day burglary, and grand larceny. Overall, Boggs’ (1965) seminal work

suggested that for some crime types, the residential population may be an inappropriate and misleading denominator.

Subsequent research conducted by Skogan (1976) compared rates of motor vehicle theft in several large American cities per 1,000 residents to rates per 1,000 vehicles. He found that while New York City ranked quite low using the traditional rate (12 motor vehicle thefts per 1,000 residents), it ranked first amongst the cities studied when the alternative rate was used (53 motor vehicle thefts per 1,000 vehicles), underscoring the importance of selecting “meaningful denominators, to analyze victimization experiences in light of the exposure of potential victims to risk” (Skogan, 1976, p. 172).

In contrast Cohen et al. (1985) suggested that concerns about the accuracy of residential population-based crime rates may be unwarranted. Cohen et al. (1985) found that when traditional and alternative rates for burglary and auto theft were compared, they were quite similar. Moreover, the traditional rates consistently provided better forecasts than the alternative rates.

More recently, researchers have calculated the ambient population measure in a variety of innovative ways, including 24-hour average population estimates from LandScan Global Population Database (Andresen, 2006a, 2011; Piza & Gilchrist, 2018), Twitter messages (Hipp et al., 2019; Kounadi et al., 2018; Malleson & Andresen, 2015a), and transportation survey data (Boivin, 2018; Felson & Boivin, 2015). These and other measures have been used to examine many different crime types, such as snatch-and-run offenses (Hanaoka, 2018), theft from person (Malleson & Andresen, 2016), stranger assaults (Boivin, 2013), violent crime (Malleson & Andresen, 2015a,b) and automotive theft

(Andresen, 2006a). Overall, the literature consistently demonstrates the value of considering this alternative denominator in crime analysis.

Andresen and Jenion (2008, 2010) found that the ambient population was more clustered in the city centre, compared to the residential population, drastically reducing measured risk in those areas in Vancouver. Similarly, Malleson and Andresen (2016) found that all four of the ambient population measures they evaluated (census workday population, geo-located Twitter messages, mobile telephone activity counts, and Population 24/7 population estimates) demonstrated clustering in London's city centre. And using LandScan Global Population Database data as an ambient population denominator, Andresen (2011) found that aggregate violent crime rates dropped in Vancouver's downtown area.

Malleson and Andresen (2015b) employed the Getis-Ord GI^* statistic and found that there was statistically significant clustering of residential population-based violent crime rates in Leeds' city centre. However, when the ambient population was accounted for, using geo-located Twitter messages, these clusters became insignificant. Later work by Malleson and Andresen (2016) identified new hotspots when significant Getis-Ord GI^* clusters of ambient population-based rates of theft from persons offenses were mapped. These clusters, where the risk of being a victim of a theft from persons offense was higher, would not have been identified using only a residential population-based crime rate.

Felson and Boivin (2015) used transportation survey data to capture the number of daily visitors in census tracts in a large Eastern Canadian city. They found that various visitor types were strongly linked to aggregate property and violent crime (Felson & Boivin, 2015). Kurland, Johnson, and Tilley (2014) compared rates of violent crime and

theft and handling offenses around a UK stadium, using LandScan Global Population Database ambient population data and match/event ticket sales to provide population at risk estimates for regular days and match/event days, respectively. By using these two ambient population measures, the authors did not rely on the residential population at all. One interesting finding was that although counts of theft and handling offenses were much higher on match and event days, compared to days when neither occurred, the rates of these offenses were significantly lower on match and event days (Kurland et al., 2014).

Within the last five years there has been a growing number of multivariate analyses conducted using ambient population measures as independent variables. In the context of commuting, Boivin (2013) used the number of workers in Montreal census tracts to estimate the ambient population. Although the residential population was found to have a significant positive effect on both domestic violence and burglaries, no such relationship existed for stranger assaults. The ambient population, however, emerged as a strong, significant predictor of the number of stranger assaults. Stults and Hasbrouck (2015) used commuting data as well to examine crime rate estimates at the municipal level, finding that daily commuting rates were a strong predictor of overall crime rates in American cities.

Both Boivin (2018) and Boivin and Felson (2018) calculated census tract visitors using transportation survey data. Boivin and Felson (2018) found that an increase in visitors was associated with both more visitors and residents being charged with a crime in that census tract. By contrast, Boivin's (2018) research suggested that the relationship between crime and visiting populations is more ambiguous. Using geographically weighted regression, Boivin (2018) found that larger visiting populations were associated with higher levels of crime. However, for some visit types, the relationship with crime was

negative. And in their assessment of routine activity theory and crime pattern theory, Hipp et al. (2019) found that their temporal ambient population measure, geolocated Twitter data, was useful in explaining crime at the city block level.

Only three studies have employed ambient population-based crime rates as dependent variables. Andresen (2006a) found that model fit was substantially higher for the models using ambient population-based crime rates. Additionally, ambient rate regression models retained more independent variables compared to residential rate models. Moreover, variables linked to both routine activities theory and social disorganization theory have consistently been associated with ambient population-based crime rates (see Andresen, 2006a, 2011).

Data and methods

The current study examines residential and ambient population-based crime rates for disaggregated property crime types in Vancouver, British Columbia at the dissemination area level. While descriptive findings and maps are presented, the study is primarily inferential. Spatial error models for both residential and ambient population-based rates are produced and compared. Model variables are informed by both social disorganization theory and routine activity theory. The ambient population measure was constructed using open source cell tower location data from OpenCellID (<https://opencellid.org/>).

Data

The dataset used in this study consists of three data sources: property crime data from Vancouver's municipal police force (VPD), Statistics Canada census data, and open source

cell tower location data from OpenCellID. The year 2016 was chosen for the VPD data to correspond with Canada's most recent census. The 2016 property crime data used in this study from the VPD consist of five crime types spatially joined to dissemination areas: mischief, theft from vehicle, theft of vehicle, theft of bicycle, and other theft. In 2016, there were 3938 counts of mischief, 8870 thefts from vehicle, 1288 thefts of vehicle, 2405 thefts of bicycle, and 5708 other thefts. These data are publicly available through the City of Vancouver's open data catalogue (<http://vancouver.ca/your-government/open-data-catalogue.aspx>).

The OpenCellID data source was used to create an ambient population measure. OpenCellID describes itself as "the world's largest collaborative community project that collects GPS positions of cell towers" (OpenCellID, 2018). Users typically join to obtain location services information on their mobile devices without relying on GPS, as well as to research cell tower coverage. As users move around, the OpenCellID software collects information about where the users are, and which cells they are connected to. This allows for estimates of the locations of the cells themselves. Hence the data represent cells in cellular networks, not the dynamic movements of OpenCellID users. Individual cells are serviced by base transceiver stations that use antennae fixed to cell towers to provide network coverage. Often, there are multiple antennae from multiple providers on a single tower. The size of the cell service area depends on a variety of factors, such as the number of users and the characteristics of the surrounding environment (e.g. topography, weather). Importantly, mobile telephone operators need to install larger number of cells in areas where there are large numbers of users, so they can be a useful proxy for the size of the ambient population. The data are cumulative, with user-identified cells being added to the

database over time. To the researchers' knowledge, no prior studies have made use of this data source to study crime. The file used in the current study was downloaded on September 14th, 2017.

The study area, the City of Vancouver, is part of the larger 'Metro Vancouver' region. Although the analysis of the crime patterns undertaken in this paper focuses solely on the City of Vancouver, there are large numbers of people who travel to the City from the surrounding municipalities, so the entire Metro Vancouver area needs to be taken into account when estimating the new ambient population. The first step in the process of estimating new ambient population is to spatially join the OpenCellID data to dissemination areas Metro Vancouver and create a count of cells per dissemination area. Within the City of Vancouver itself, 19215 unique cells were identified with dissemination area cell counts ranging from zero to 732. Not surprisingly, the dissemination area with the highest cell count was located in the City of Vancouver's downtown core. Cells also clustered at major population centers and along transportation corridors.

Following the calculation of cell counts per dissemination area, the residential population of Metro Vancouver needs to be redistributed from where people *live* to where they are likely to undertake their day-to-day activities. The residential population, as estimated in the 2016 Census, was proportionately redistributed based on the number of cells in each dissemination area throughout the Metro Vancouver region. Hence we assume that the ambient population estimated here is a daily average of the size of the population in each dissemination area. For the purposes of this study, Metro Vancouver was considered a closed system, so we disregard the possible flows into the City of Vancouver from outside the metropolitan region. In cases where dissemination areas had zero cells,

the average calculated ambient population from the nearest spatial neighbours was used (Queen's contiguity 1). At approximately 800,000 persons, the calculated ambient population based on OpenCellID data for the City of Vancouver is 27% greater than its residential population.

<Insert Figure 1 About Here>

Figure 1 highlights the differences in the spatial patterning of Vancouver's residential and ambient populations. While the residential population of Vancouver's dissemination areas ranges from 68 to 8778, the ambient population ranges from 41 to 29971. In terms of percent change, the difference between the residential population and the ambient population in Vancouver's dissemination areas ranged from approximately 95 percent all the way to 2266 percent. These figures depict ambient population clustering in Vancouver's downtown core. As discussed in the literature review, using the residential population as a crime rate denominator in this area could lead to spuriously high crime rates.

Three limitations of this ambient population measure should be noted. First, the locations of the cells are averaged based on multiple measurements from OpenCellID users, meaning that their recorded locations may differ slightly from their actual locations. As the cell locations are aggregated to dissemination areas anyway, these minor errors are unlikely to impact on the analysis. Second, and somewhat obviously, the data from OpenCellID is user-generated. As noted by Malleson and Andresen (2015a) in their study that estimated the ambient population using geo-located Twitter messages, there may be omissions and biases. For instance, homeless and poorer populations may have lower rates of mobile service subscription. These populations may be under-represented in the current

study's ambient population measure. Lastly, because the data are cumulative, seasonal and event-driven population changes cannot be accounted for.

The 2016 Statistics Canada census data used in this study are measured at the dissemination area level, the smallest census unit in Canada with sociodemographic and socioeconomic data, typically having a population of 400-700 persons, similar to the US census block group. In 2016 there were 991 dissemination areas in Vancouver. Of these 991 dissemination areas, 13 were excluded from this analysis because, for confidentiality reasons (due to low residential population counts). Variables were chosen based on their relevance to social disorganization theory and routine activities theory—see Andresen (2020) for a detailed discussion of these theories and their empirical support. As mentioned above, population/ethnic composition is a key construct in social disorganization theory (Shaw & McKay, 1942). Percentages of Aboriginals, visible minorities, immigrants, and ethnic heterogeneity were chosen to reflect this construct. The ethnic heterogeneity variable was calculated from census data on ethnic origins. Using this data, scores on the Blau (1977) index were generated. A score of zero indicates no mix of ethnic groups (i.e. ethnic homogeneity), whereas a score of one hundred indicates an even ethnic mix (i.e. perfect ethnic heterogeneity). Percentages of recent immigrants, people who moved into the dissemination area within the last year, and rented households were used to capture population turnover and residential mobility (Sampson & Groves, 1989; Shaw & McKay, 1942). Rented households also have relevance when it comes to the routine activity theory concept of guardianship; renters are expected to engage in more activities away from home (Andresen, 2006b). To measure economic status, median income along with the percentages of unemployment, government assistance, low income designation,

subsidized housing, housing under major repair, and post secondary education levels were used. Family disruption (Sampson & Groves, 1989) was measured using the percentage of lone parents. Lastly, the number of young males (aged 15-24) and single people were used because of their increased likelihood of victimization under the routine activity framework (Cohen & Felson, 1979; Kennedy & Forde, 1990). Young males have also been associated with increased criminal activity (Hirschi & Gottfredson, 1983).

Methods

To permit comparison, crime rates with both residential and ambient population denominators were constructed. These ten rate variables were then used as dependent variables in ten separate regression models. Spatial error models were identified through statistical testing and Queen's contiguity orders for the models were determined with Moran's I significance testing of the error residuals. Due to the presence of heteroskedasticity in the residuals we considered spatial heteroskedasticity and autocorrelation (KP-HET) proposed by Kelejian and Prucha (2010) and used GeoDaSpace (<https://spatial.uchicago.edu/software>) to estimate GMM spatial error models with KP HET standard errors. Statistical models with all 17 independent variables (full models) were produced for both the resident and ambient population-based rates for each of the five crime types. In terms of specification for final statistical models, the least significant variable were removed first, then the regressions were re-run. This process was repeated until all the remaining variables were significant at the $p < 0.10$ level (Andresen, 2006a). This was done to minimize the chances of omitted variable bias in the final statistical models. These full and final statistical models permit comparison between residential and

ambient population-based crime rates on model fit, variable retention, and significant relationships; this comparison shows that the full and final models are qualitatively similar and that multicollinearity and omitted variable bias are not a concern.

Results

Descriptive results

Table 1 presents descriptive statistics for the dependent variables used in the current study. The ranges, means, and standard deviations for ambient population-based rates of mischief, theft of vehicle, theft from vehicle, and theft of bicycle are consistently greater, compared to their residential counterparts. The reverse is true for the other theft crime type. Compared to residential population-based property crime rates, nearly all of the crime rates demonstrate lower ambient rates in Vancouver's north-central downtown area. Because of the larger ambient population in these dissemination areas, crime rates using this denominator will necessarily be lower. This is most notable in Figures 2 and 3 that show the resident- and ambient-based rates of theft from vehicle. The resident rate clearly shows the theoretically expected concentrations of crime in or close to Vancouver's downtown area. However, the ambient-based rate shows concentrations of theft from vehicle that are spread throughout the city—very different spatial patterns of crime with different populations at risk.

<Insert Table 1 and Figures 2 and 3 About Here>

Descriptive statistics for all independent variables used in the current study are presented in Table 2, and Spearman's correlations for the independent variables are presented in Table 3. Only two of the significant relationships are above the commonly

used 0.8 threshold for multicollinearity. The relationship between visible minorities and immigrants ($\rho = 0.906, p < 0.01$) is hardly surprising; in Vancouver/Canada immigrants are often also visible minorities (Ley & Smith, 2000). It would therefore make sense that dissemination areas with greater percentages of immigrants would have greater percentages of visible minorities, and vice versa. Those receiving government assistance and median income are also highly correlated ($\rho = -0.834, p < 0.01$). This relationship is intuitive: as the percentage of those receiving government assistance in a dissemination area increases, median incomes decrease. All four of the above variables were kept in this analysis to avoid omitted variable bias.

<Insert Tables 2 and 3 About Here>

Inferential results

The Pseudo R2 values for the full and final models for the residential population-based mischief rate are 0.217 and 0.174, respectively—see Table 4. The percentage of single persons in Vancouver dissemination areas was identified as the biggest predictor of mischief across both models (Full residential: $\beta = 0.265, p < 0.01$; Final residential: $\beta = 0.28, p < 0.01$). The percentage of rented households is negatively associated with rates mischief across both models (Full residential: $\beta = -0.051, p < 0.05$; Final residential: $\beta = -0.028, p < 0.1$). The percentages of those receiving government assistance ($\beta = -0.155, p < 0.05$) and those with post-secondary education ($\beta = -0.067, p < 0.1$) are both negatively associated with rates of mischief in the final residential model.

For the full ambient population model of mischief, the Pseudo R2 value is 0.116, while the final model has a value of 0.105. Six variables have significant relationships with

ambient population-based rates of mischief, and these variables were all retained in the final models. The percentage of aboriginals in a dissemination area emerged as the most important predictor of mischief (Full ambient: $\beta = 0.719$, $p < 0.01$; Final ambient: $\beta = 0.87$, $p < 0.01$). Increased percentages of those receiving government assistance (Full ambient: $\beta = 0.401$, $p < 0.05$; Final ambient: $\beta = 0.37$, $p < 0.05$), low income designation (Full ambient: $\beta = 0.23$, $p < 0.05$; Final ambient: $\beta = 0.165$, $p < 0.1$), and those with post-secondary education (Full ambient: $\beta = 0.141$, $p < 0.05$; Final ambient: $\beta = 0.151$, $p < 0.01$) are all associated with higher ambient population-based rates of mischief. The percentages of both residents who moved into a dissemination area within the last year (Full ambient: $\beta = -0.184$, $p < 0.01$; Final ambient: $\beta = -0.187$, $p < 0.05$) and lone parents (Full ambient: $\beta = -0.224$, $p < 0.05$; Final ambient: $\beta = -0.245$, $p < 0.05$) are negatively associated with mischief.

<Insert Table 4 About Here>

The full and final models for residential population-based rates of theft from vehicle have respective Pseudo R2 values of 0.168 and 0.145—see Table 5. As percentages of visible minorities (Full residential: $\beta = -0.164$, $p < 0.01$; Final residential: $\beta = -0.15$, $p < 0.01$) and lone parents (Full residential: $\beta = -0.181$, $p < 0.1$; Final residential: $\beta = -0.238$, $p < 0.05$) increase, rates of theft from vehicle decrease. As the median income of a dissemination area increases, rates of theft from vehicle also decrease (Full residential: $\beta = -0.206$, $p < 0.05$; Final residential: $\beta = -0.229$, $p < 0.05$). In the full model, the percentage of subsidized housing has a positive relationship with residential population-based rates of theft from vehicle ($\beta = 0.076$, $p < 0.1$), while the percentage of those receiving government assistance has a negative one ($\beta = -0.264$, $p < 0.05$). Lastly, houses under major repair emerged as a significant predictor of theft from vehicle in the final residential model ($\beta = 0.139$, $p < 0.1$).

In other words, as the percentage of houses under major repair in a dissemination area increases, so does the residential population-based rate of theft from vehicle.

The Pseudo R2 values for the full and final ambient population-based models of theft from vehicle are 0.075 and 0.054, respectively. All four of the significant associations were maintained across the full and final models. The percentages of those receiving government assistance (Full ambient: $\beta = 0.603$, $p < 0.1$; Final ambient: $\beta = 1.166$, $p < 0.01$) and those with post-secondary education (Full ambient: $\beta = 0.258$, $p < 0.1$; Final ambient: $\beta = 0.385$, $p < 0.01$) are associated with increased ambient population-based rates of theft from vehicle. Those receiving government assistance is also the most important predictor for both the full and final ambient models. As the percentages of single persons (Full ambient: $\beta = -0.532$, $p < 0.01$; Final ambient: $\beta = -0.406$, $p < 0.05$) and lone parents (Full ambient: $\beta = -0.48$, $p < 0.05$; Final ambient: $\beta = -0.374$, $p < 0.1$) increase, rates of theft from vehicle decrease.

<Insert Table 5 About Here>

The Pseudo R2 values for the full and final models for the residential population-based rate of theft of vehicle are 0.105 and 0.091, respectively—see Table 6. The biggest predictor across both models is single persons (Full residential: $\beta = 0.031$, $p < 0.05$; Final residential: $\beta = 0.042$, $p < 0.01$). This finding means that as the percentage of single persons in a dissemination area increases, so does the rate of theft of vehicle. The percentages of immigrants (Full residential: $\beta = -0.024$, $p < 0.1$; Final residential: $\beta = -0.024$, $p < 0.01$) and those with post-secondary education (Full residential: $\beta = -0.025$, $p < 0.05$; Final residential: $\beta = -0.026$, $p < 0.01$) have negative relationships with the rate of theft of vehicle. The percentage of houses under major repair in a dissemination area became

significant in the final model for theft of vehicle ($\beta = 0.031, p < 0.1$). Ethnic heterogeneity also emerged as a significant negative predictor of residential population-based rates of theft of vehicle in the final model ($\beta = -0.012, p < 0.05$).

For the ambient population-based rates of theft of vehicle, the Pseudo R2 value for the full model is 0.057, while that of the final model is 0.045. For this crime rate, all three significant variables were retained in the final model. The largest predictor of theft of vehicle was found to be the percentage of aboriginals (Full ambient: $\beta = 0.249, p < 0.05$; Final ambient: $\beta = 0.261, p < 0.01$). The percentage of people on government assistance (Full ambient: $\beta = 0.181, p < 0.05$) is also associated with increased rates of theft of vehicle. Dissemination areas with a greater percentage of lone parents are associated with lower ambient population-based rates of theft of vehicle (Full ambient: $\beta = -0.094, p < 0.1$; Final ambient: $\beta = -0.115, p < 0.05$).

<Insert Table 6 About Here>

The respective Pseudo R2 values for the full and final residential population-based rate models for theft of bicycle are 0.196 and 0.173—see Table 7. The biggest predictor across both models is those receiving government assistance (Full model: $\beta = -0.203, p < 0.01$; Final model: $\beta = -0.217, p < 0.01$). As the percentage of residents receiving government assistance in a dissemination area increases the residential population-based rate of theft of bicycle decreases. The percentage of immigrants is also negatively associated with the rate of theft of bicycle (Full residential: $\beta = -0.084, p < 0.05$; Final residential: $\beta = -0.07, p < 0.01$). For both models, the percentages of single persons (Full residential: $\beta = 0.156, p < 0.01$; Final residential: $\beta = 0.152, p < 0.01$) and low-income designation (Full residential: $\beta = 0.108, p < 0.1$; Final residential: $\beta = 0.115, p < 0.05$) are

positively associated with theft of bicycle. When the percentages of each of these variables increases, so does the residential population-based rate of theft of bicycle in Vancouver dissemination areas.

For the ambient population-based models of theft of bicycle in Vancouver, the full model has a Pseudo R2 value of 0.121, while the final model has a value of 0.11. This time, the percentage of males aged 15-24 emerged as the largest significant predictor across both full and final models (Full ambient: $\beta = -0.664$, $p < 0.05$; Final ambient: $\beta = -0.593$, $p < 0.01$). Dissemination areas with a greater percentage of young males have lower ambient population-based rates of theft of bicycle. Ethnic heterogeneity (Full model: $\beta = -0.044$, $p < 0.1$; Final model: $\beta = -0.052$, $p < 0.05$) and visible minorities (Full model: $\beta = -0.059$, $p < 0.1$; Final model: $\beta = -0.099$, $p < 0.01$) were also found to be negatively associated with theft of bicycle. Low income is a positive predictor across both models (Full model: $\beta = 0.156$, $p < 0.05$; Final model: $\beta = 0.179$; $p < 0.01$). Finally, the percentage of those with post-secondary education are associated with increased theft of bicycle ($\beta = 0.062$, $p < 0.1$), but only in the full ambient model.

<Insert Table 7 About Here>

For the final crime type examined, other theft, the full and final residential population-based models for other theft have Pseudo R2 values of 0.051 and 0.041, respectively—see Table 8. There is not consistency across the full and final models in terms of the most important predictor. Both ethnic heterogeneity (Full residential: $\beta = 0.138$, $p < 0.1$; Final residential: $\beta = 0.118$, $p < 0.1$) and the percentage of single persons (Full residential: $\beta = 0.61$, $p < 0.01$; Final residential: $\beta = 0.671$, $p < 0.01$) are positively associated with rates of other theft. In only the full residential model, median income has a

positive relationship with other theft rates ($\beta = 0.322, p < 0.1$), while the unemployment percentage has a negative one ($\beta = -0.663, p < 0.1$). In the final residential model, both the percentages of immigrants ($\beta = 0.202, p < 0.1$) and those receiving government assistance ($\beta = -0.527, p < 0.05$) emerged as significant predictors.

The full ambient population-based model for other theft has a Pseudo R2 value of 0.047, while the final model has a value of 0.04. Compared to the residential population-based models, the ambient ones retain more variables. Across both models, the percentage of young males was found to be the biggest predictor of other theft (Full ambient: $\beta = -0.435, p < 0.05$; Final ambient: $\beta = -0.549, p < 0.05$). This finding means that as the percentage of young males in a dissemination area increases, the ambient population-based rates of other theft actually decrease. This somewhat counterintuitive finding will be discussed more in the following section. Government assistance also has a negative relationship with ambient population-based rates of other theft (Full ambient: $\beta = -0.264, p < 0.05$; Final ambient: $\beta = -0.247, p < 0.01$). The percentages of recent immigrants (Full ambient: $\beta = 0.227, p < 0.1$; Final ambient: $\beta = 0.296, p < 0.05$) and single persons (Full ambient: $\beta = 0.13, p < 0.1$; Final ambient: $\beta = 0.188, p < 0.01$) have positive relationships with the dependent variable. In the final model, the percentage of lone parents is a significant predictor of increased ambient population-based rates of other theft ($\beta = 0.186, p < 0.05$).

<Insert Table 8 About Here>

Discussion and conclusions

The differences in spatial patterns of rates of mischief, theft from vehicle, theft of vehicle, theft of bicycle, and other theft are often striking, depending on the population denominator used. The finding that hotspots in Vancouver's downtown area decrease in intensity when the ambient population is used speaks to Boggs' (1965) assertion regarding spuriously high crime rates in central business districts. When a more appropriate population at risk is used, the risk of being a victim of property crime is not substantially higher in downtown Vancouver dissemination areas.

New, and often more dispersed, clusters of high crime rate dissemination areas were identified in the current study when maps of ambient population-based rates were compared to residential ones. These results echo the work of Malleson and Andresen (2015b, 2016), who also identified new statistically significant clusters of aggregate violent crime and theft from persons offenses when an ambient population measure was used. Overall, these findings underscore the importance of considering the population at risk. Because such a different picture of environmental risk is painted when the ambient population is used as the crime rate denominator, these findings seriously bring into question the near-ubiquitous use of the residential population.

The multivariate results are the primary focus of the current study. Figure 4 shows significant relationships, variable retention, and Pseudo R2 values between regression models. A quick glance at Figure 4 reveals important differences between regression models for disaggregated property crime rates that use either residential or ambient population denominators. The current study's finding that Pseudo R2 values are lower for full and final models for ambient population-based crime rates, compared to their

residential counterparts, is somewhat unexpected. Previous studies have consistently found the opposite (Andresen, 2006a, 2011). Only for aggregate violent crime did Andresen (2006a) find a higher Pseudo R2 value for the ambient model. The opposite trend of the current study to prior works may be explained by the differences between the studies detailed above. The generally low Pseudo R2 values suggest that there is more to explain in the spatial patterns of crime at the dissemination area level than the theoretically-informed census variables permit.

<Insert Figure 4 About Here>

In terms of variable retention across full and final models, there is either more or equal retention for ambient population-based rates. Final models for ambient population-based rates also typically have either a greater or equivalent number of significant variables, compared to final residential models. The only exception is theft of vehicle. These findings show the limited value of using Pseudo R2 values for model assessment. For the three prior inferential studies that used the ambient population as a crime rate denominator, there was no consistent pattern in terms of the number of significant variables.

Only a handful of significant variables are consistent across final residential and ambient models in the direction of their relationship with the particular crime type. The percentage of lone parents in a dissemination area has a negative relationship with theft from vehicle, regardless of the population denominator used. Whichever population at risk is accounted for, greater percentages of lone parents are associated with a decrease in this crime type. This relationship may be a question of suitable targets; lone parents may be less able to afford a vehicle to be broken into.

Low income has a consistent positive relationship with theft of bicycle for both residential and ambient population-based rate models. This finding may speak to social disorganization processes regarding low socioeconomic status and crime (Sampson & Groves, 1989). For other theft, the percentage of single persons in a dissemination area is a positive predictor for both residential and ambient models. The routine activities of this population segment may bring them out of the home more, which may create more opportunities for victimization (Cohen & Felson, 1979). Finally, government assistance is negatively associated with both residential and ambient population-based rates of other theft.

Immediately apparent from Figure 4 are the differences between the final residential and ambient population-based rate models. Many variables that are significant predictors in one model are insignificant in the other. For example, low income designation is a positive predictor of ambient population-based rates of mischief. This finding speaks to social disorganization theory, and the link between low socioeconomic status and crime (Sampson & Groves, 1989). However, the same relationship does not hold up when the residential population is used. From a theoretical perspective, the finding is somewhat troubling. If a relationship predicted by social disorganization theory holds only when the ambient population is used, it brings into question the exclusive use of the residential population as a crime rate denominator for theory testing.

In the context of routine activity theory, the relationships between the percentage of single persons in a dissemination area and rates of mischief, theft of vehicle, and theft of bicycle are quite interesting. For residential population-based rates of these three crime types, the percentage of single persons is a significant positive predictor. This relationship

is consistent with Cohen and Felson's (1979) findings. Single people's routine activities take them out of the home more often and put them at greater risk of criminal victimization (Cohen & Felson, 1979). Yet, when the ambient population is used all three of these relationships become insignificant. Both this and the above finding highlight the impact an alternative denominator can have on theoretically-predicted relationships.

The differences between final residential and ambient population-based rate models are also important when it comes to policy-relevant variables. While ethnic heterogeneity or the percentage of lone parents in a dissemination area cannot (reasonably) be controlled, policies enacted by various levels of government on subsidized housing, for instance, can affect crime rates. In the current study, percentages of post-secondary education are negatively associated with residential population-based rates of theft of vehicle. Policymakers might think that improving access to post-secondary education could have long-term effects on rates of theft of vehicle. When the ambient population is used, however, this relationship disappears. If this measure provides a more accurate indication of environmental risk than the residential population, policies enacted to increase post-secondary education may be ineffective. A similar trend exists for those receiving government assistance and rates of theft of bicycle. When the residential population is used, the relationship is negative, but becomes insignificant in the ambient model. Policy decisions depend on accurate assessments of the relationships between crime risk and sociodemographic and socioeconomic indicators. These findings suggest that alternative population measures should be considered alongside the residential population, when conducting research relevant to policy.

While most of the differences between final residential and ambient population-based rate models involve variables falling in and out of significance, there are two instances of relationships switching direction between final models for residential and ambient population-based crime rates. The percentage of those receiving government assistance has a negative relationship with residential population-based rates of mischief, yet the relationship is positive for the ambient rate. This finding means that when the residential population of a dissemination area is controlled for, increased percentages of those receiving government assistance are associated with lower rates of mischief, and vice versa. However, when it is the ambient population that is controlled for, both percentages of those receiving government assistance and rates of mischief vary together. Interestingly, when the ambient population is used, the relationship conforms to social disorganization expectations regarding low socioeconomic status. Yet the negative relationship in the residential model may speak more to routine activity theory. In dissemination areas with lower percentages of those receiving government assistance, there may be more suitable targets for mischief.

A similar trend exists for postsecondary education and mischief. Postsecondary education is negatively associated with residential population-based rates of mischief, but the relationship is positive when the ambient population is controlled for instead. These findings are particularly important, because they demonstrate the impact the use of a theoretically-informed alternative denominator can have on results. When the number of people that visit an area are considered, as opposed to the number of people who sleep in that area, significant relationships can switch direction. Worth noting, is that there was no

switching of signs for socioeconomic or sociodemographic variables in the studies conducted by Andresen (2006a, 2011).

From a policy perspective, these findings are perhaps even more worrisome than variables going in and out of significance between final residential and ambient population-based crime rate models. Crime reduction policies are often informed by relationships between residential population-based crime rates and sociodemographic and socioeconomic indicators. If more accurate population measures (i.e. the ambient population) suggest that these same relationships are in the opposite direction, these policies could potentially increase crime. Taken together, the findings from this study indicate that the use of a theoretically-informed crime rate denominator impacts results in a substantial way. There are important differences in spatial patterns, Pseudo R² values, variable retention, and trends in significant relationships between crime rates using either residential or ambient population denominators. As discussed, there are differences between the current study and the work of Andresen (2006a, 2011) in both design and results. Nevertheless, the overall story told is the same. Clearly, the question of the most appropriate crime rate denominator is not just an obscure measurement issue to be acknowledged in passing; the population at risk matters.

The ambient population measure is not without limitations, the most noteworthy being omissions and biases related to OpenCellID users themselves. Obviously, not everyone uses OpenCellID; the cell tower location data reflect the movements of those who do. It would also not be a stretch to suggest that OpenCellID users are probably younger than the average mobile phone user, given their decision to use such an app. Still, so long as

some OpenCellID users pass near a cell at some point in the preceding years, the cell will be added to the database.

There is considerable scope to improve the accuracy of the ambient population estimates created here through the incorporation of additional data sources. For example, the OpenCellID data could be supplemented with information about the: density of buildings that are associated with large ambient populations (e.g. offices, transportation hubs, etc.); volume of activity on spatially-located social media services such as Twitter; results of travel or time-use surveys; ambient light levels; etc. Furthermore, by linking with temporally dynamic data sources or those that contain socio-demographic information, it might be possible to move towards a more nuanced, granular, representation of the ambient population. Ultimately however, the spatial patterns of cell density used here do reflect local knowledge about population centers and transportation corridors in Metro Vancouver and hence the ambient population measure constructed from this data source likely provides a better estimation of the population at risk than the residential population.

Two other limitations relate to the use of census data. First, it has been suggested that census data does not directly measure social disorganization constructs (Andresen, 2020). Rather, self-report data is necessary to adequately capture mediating factors such as sparse local friendship networks (see Sampson & Groves, 1989; Lowenkamp et al., 2003). A similar argument could be made for routine activity theory. Variables such as median income only act as proxies for concepts like the number of suitable targets in an area.

A second limitation of census data concerns their link to residents of spatial units like dissemination areas. Some of the results from the current study were surprising, such as the negative relationship between the percentage of males aged 15-24 and ambient

population-based rates of theft of bicycle and other theft. Past research has consistently linked young males with increased crime rates (Hirschi & Gottfredson, 1983). This finding may be a question of where people live versus where their routine activities take them. All census variables correspond to residential, not ambient populations. This means that even when ambient crime rates are used in regression models to capture exposure to risk, the independent variables are still based on the residential population. There is no practical solution to this problem, but it should be acknowledged. This issue may explain the lower Pseudo R² values for ambient population-based rate models as a result of omitted variable bias. The 'right' independent variables that correspond to sociodemographic and socioeconomic indicators for ambient populations are unavailable. Regardless, we argue that this is a lesser concern than properly measuring exposure in the dependent variable. For example, if a high socioeconomic status (SES) neighbourhood "loses" half of its residential population during the day for work and has the same number of lower SES people entering the neighbourhood for work, the character of the neighbourhood will still be high SES: these people will work in the homes of the high SES people and in the expensive local stores in the area as well such that the characteristics of the neighbourhood do not change as much as exposure to risk when people come and go in their neighbourhoods.

Despite the limitations detailed above, the results from this study clearly demonstrate the importance and value of considering the ambient population in crime analysis. While the current research cannot say definitively whether this particular measure of the ambient population provides a better estimation of the population at risk than the residential population, it is clear that it impacts both spatial patterns and

regression results substantially. Future studies should make use of the ambient population alongside the residential population, as the data are now easier than ever to obtain (Andresen, 2006a). This and other ambient population measures should be applied to different settings, at different spatial scales, and with disaggregated crime data. Regarding the aggregation of crime data, many of the studies discussed in this paper employed aggregate measures of crime (see Andresen, 2006a, 2011; Kurland et al., 2015). It is entirely possible that important trends are being masked when various crime types are aggregated into a single measure. Overall, more widespread use of the ambient population is recommended, particularly in a multivariate context.

References

- Andresen, M. A. (2006a). Crime measures and the spatial analysis of criminal activity. *The British Journal of Criminology*, 46(2), 258-285.
- Andresen, M. A. (2006b). A spatial analysis of crime in Vancouver, British Columbia: A synthesis of social disorganization and routine activity theory. *The Canadian Geographer*, 50(4), 487-502.
- Andresen, M. A. (2011). The ambient population and crime analysis. *The Professional Geographer*, 63(2), 193-212.
- Andresen, M. A. (2020). *Environmental criminology: Evolution theory and practice (2nd ed.)*. New York, NY: Routledge.
- Andresen, M. A., & Jenion, G. W. (2008). Crime prevention and the science of where people are. *Criminal Justice Policy Review*, 19(2), 164-180.
- Andresen, M. A., & Jenion, G. W. (2010). Ambient populations and the calculation of crime rates and risk. *Security Journal*, 23(2), 114-133.
- Andresen, M. A., & Malleson, N. (2013). Spatial heterogeneity in crime analysis. In M. Leitner (Ed.), *Crime modeling and mapping using geospatial technologies* (pp. 3-23). New York, NY: Springer.
- Blau, P. (1977). *Inequality and heterogeneity*. New York, NY: The Free Press.
- Boggs, S. L. (1965). Urban crime patterns. *American Sociological Review*, 30(6), 899-908.
- Boivin, R. (2013). On the use of crime rates. *Canadian Journal of Criminology and Criminal Justice*, 55(2), 263-277.
- Boivin, R. (2018). Routine activity, population(s) and crime: Spatial heterogeneity and conflicting propositions about the neighborhood crime-population link. *Applied*

Geography, 95, 79-87.

Boivin, R., & Felson, M. (2018). Crimes by visitors versus crimes by residents: The influence of visitor inflows. *Journal of Quantitative Criminology*, 34(2), 465-480.

Cohen L. E., & Felson M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 44(4), 588-608.

Cohen, L. E., Kaufman, R. L., & Gottfredson, M. R. (1985). Risk-based crime statistics: A forecasting comparison for burglary and auto theft. *Journal of Criminal Justice*, 13(5), 445-457.

Felson, M., & Boivin, R. (2015). Daily crime flows within a city. *Crime Science*, 4, Article 31.

Hanaoka, K. (2018). New insights on relationships between street crimes and ambient population: Use of hourly population data estimated from mobile phone users' locations. *Environment and Planning B: Urban Analytics and City Science*, 45(2), 295-311.

Harries, K. D. (1991). Alternative denominators in conventional crime rates. In P. J. Brantingham & P. L. Brantingham (Eds.). *Environmental criminology* (pp. 147-165). Prospect Heights, IL: Waveland Press, Inc.

Hipp, J. R., Bates, C., Lichman, M., & Smyth, P. (2019). Using social media to measure temporal ambient population: Does it help explain local crime rates? *Justice Quarterly*, 36(4), 718 – 748.

Hirschi, T., & Gottfredson, M. (1983). Age and the explanation of crime. *American Journal of Sociology*, 89(3), 552-584.

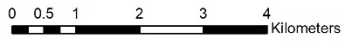
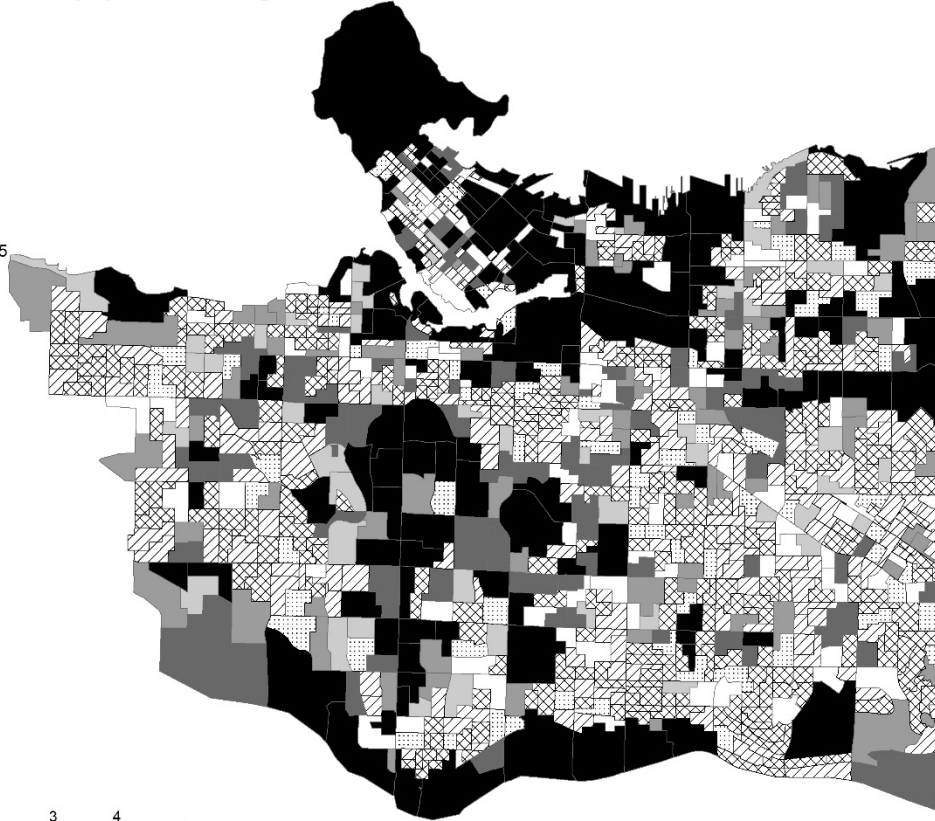
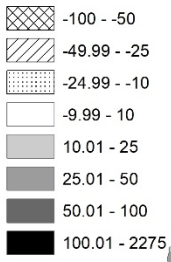
Kennedy, L. W., & Forde, D. R. (1990). Routine activities and crime: An analysis of

- victimization in Canada. *Criminology*, 28(1), 137-152.
- Kounadi, O., Ristea, A., Leitner, M., & Langford, C. (2018). Population at risk: Using areal interpolation and Twitter messages to create population models for burglaries and robberies. *Cartography and Geographic Information Science*, 45(3), 205-220.
- Kurland, J., Johnson, S. D., & Tilley, N. (2014). Offenses around stadiums: A natural experiment on crime attraction and generation. *Journal of Research in Crime and Delinquency*, 51(5), 5- 28.
- Ley, D., & Smith, H. (2000). Relations between deprivation and immigrant groups in large Canadian cities. *Urban Studies*, 37(1), 37-62.
- Lottier, S. (1938). Distribution of criminal offenses in sectional regions. *Journal of Criminal Law and Criminology*, 29(3), 329-344.
- Lowenkamp, C. T., Cullen, F. T., and Pratt, T. C. (2003). Replicating Sampson and Groves's test of social disorganization theory: Revisiting a criminological classic. *Journal of Research in Crime and Delinquency*, 40 (4), 351-373.
- Malleson, N., & Andresen, M. A. (2015a). Spatio-temporal crime hotspots and the ambient population. *Crime Science*, 4(1), Article 10.
- Malleson, N., & Andresen, M. A. (2015b). The impact of using social media data in crime rate calculations: Shifting hot spots and changing spatial patterns. *Cartography and Geographic Information Science*, 42(2).
- Malleson, N., & Andresen, M. A. (2016). Exploring the impact of ambient population measures on London crime hotspots. *Journal of Criminal Justice*, 46, 52-63.
- OpenCellID. (2018). What is OpenCellID? Retrieved from http://wiki.opencellid.org/wiki/What_is_OpenCellID

- Piza, E. L., & Gilchrist, A. M. (2018). Measuring the effect heterogeneity of police enforcement actions across spatial contexts. *Journal of Criminal Justice*, 54, 76-87.
- Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, 94(4), 774-802.
- Shaw, C. R., & McKay, H. D. (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.
- Skogan, W. G. (1976). Victimization surveys and criminal justice planning. *University of Cincinnati Law Review*, 45(2), 167-206.
- Stults, B. J., & Hasbrouck, M. (2015). The effect of commuting on city-level crime rates. *Journal of Quantitative Criminology*, 31(2), 331-350.

Figure 1. Ambient to residential population change, percent

Ambient-residential population change, %



N

Figure 2. Resident-based theft from vehicle rate

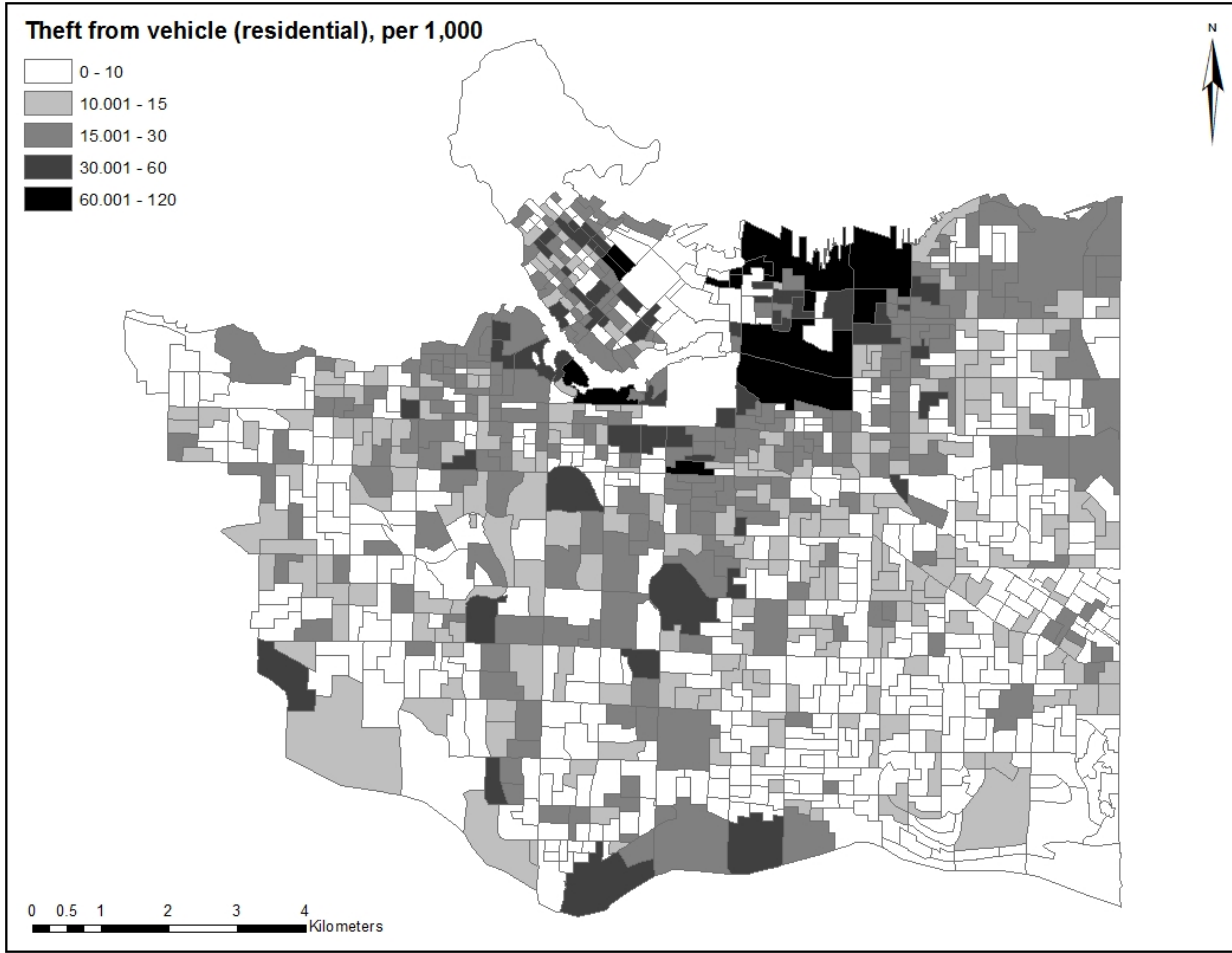


Figure 3. Ambient-based theft from vehicle rate

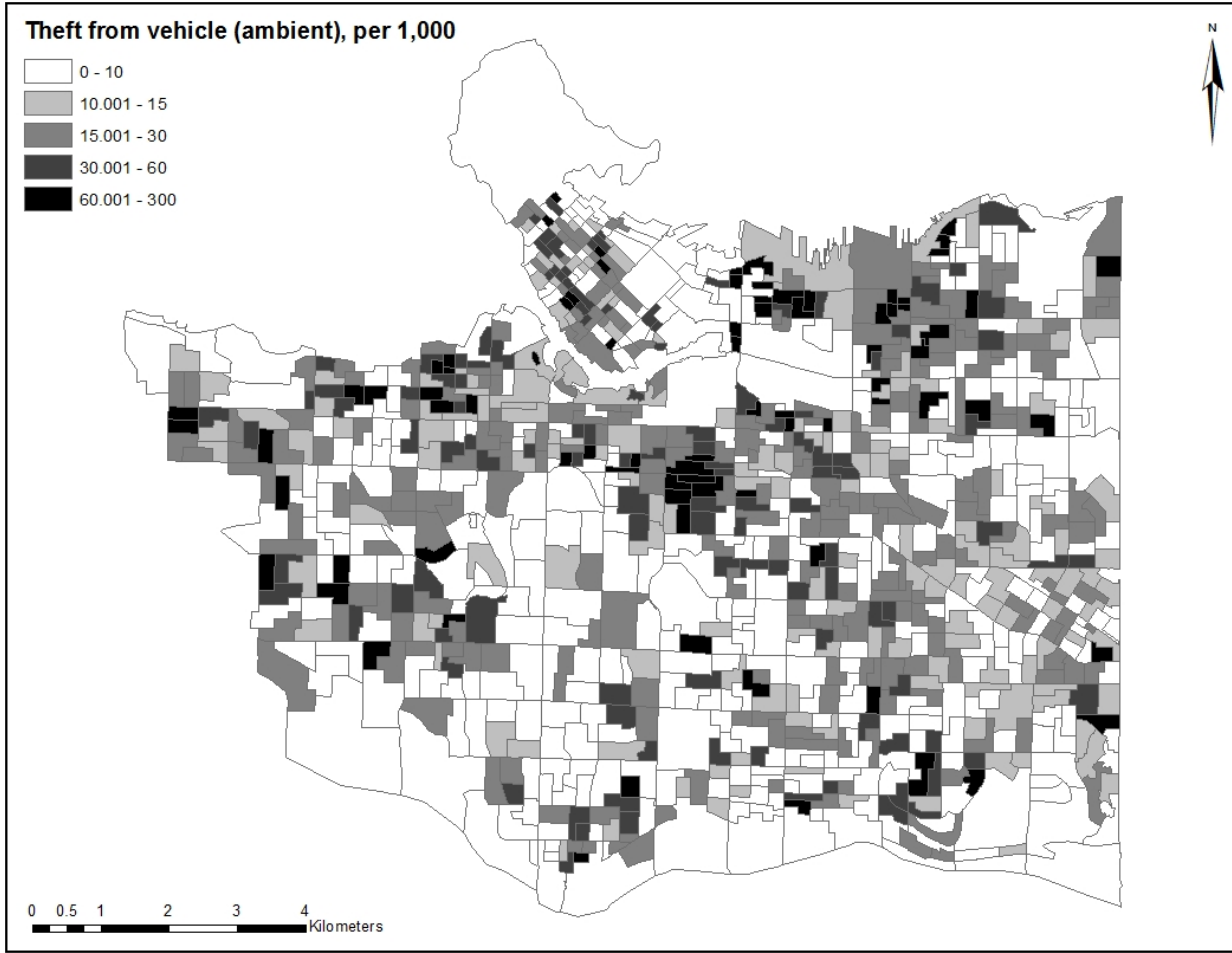


Figure 4. Regression summary

	Mischief				Theft from Vehicle				Theft of Vehicle				Theft of Bicycle				Other Theft			
	FullR	FinalR	FullA	FinalA	FullR	FinalR	FullA	FinalA	FullR	FinalR	FullA	FinalA	FullR	FinalR	FullA	FinalA	FullR	FinalR	FullA	FinalA
Aboriginal (%)			■	■							■	■								
Ethnic heterogeneity										■					■	■		■		
Visible minorities (%)					■	■									■	■				
Immigrants (%)									■	■			■	■				■		
Recent immigrants (%)																			■	■
Moved within 1 year (%)			■	■																
Single persons (%)	■	■					■	■	■	■			■	■			■	■	■	■
Lone parents (%)			■	■	■	■	■	■			■	■							■	■
Males, aged 15-24 (%)														■	■			■	■	
Unemployed (%)																	■			
Receiving gov't assistance (%)		■	■	■	■		■	■			■	■	■	■				■	■	■
Low income designation (%)			■	■									■	■	■	■				
Median income (000s, CAD)					■	■											■			
Subsidized housing (%)					■															
Houses under major repair (%)						■				■										
Rented households (%)	■	■																		
Post-secondary education (%)		■	■	■			■	■	■	■					■					
# Sig. variables	2	4	6	6	5	4	4	4	3	5	3	3	4	4	5	4	4	4	4	5
# Sig. variables retained	2		6		3		4		3		3		4		4		3		4	
Pseudo R ²	0.217	0.174	0.116	0.105	0.168	0.145	0.075	0.054	0.105	0.091	0.057	0.045	0.196	0.173	0.121	0.11	0.051	0.041	0.047	0.04

■ Significant positive relationship
 ■ Significant negative relationship

Table 1. Descriptive statistics for dependent variables

	Minimum	Maximum	Mean	Standard Deviation
Mischief (residential)	0	117.55	5.815	8.093
Mischief (ambient)	0	122.13	8.952	14.312
Theft from vehicle (residential)	0	111.111	14.002	12.734
Theft from vehicle (ambient)	0	293.112	24.174	34.498
Theft of vehicle (residential)	0	15.564	2.103	2.38
Theft of vehicle (ambient)	0	73.278	3.91	7.536
Theft of bicycle (residential)	0	96.026	3.306	6.596
Theft of bicycle (ambient)	0	146.556	4.73	11.218
Other theft (residential)	0	471.287	6.338	32.615
Other theft (ambient)	0	219.814	3.644	13.413

Note: All rates are per 1,000 persons, n = 978

Table 2. Descriptive statistics for independent variables

	Minimum	Maximum	Mean	Standard Deviation
Aboriginal (%)	0	40.404	2.259	3.561
Ethnic heterogeneity	0	93.378	39.551	16.367
Visible minorities (%)	7.407	100	50.956	25.355
Immigrants (%)	9.532	88.034	41.85	14.934
Recent immigrants (%)	0	27.933	5.78	4.037
Moved within 1 year (%)	0	62.793	16.225	7.71
Single persons (%)	23.81	90.121	43.411	9.431
Lone parents (%)	0	62.791	16.159	7.085
Males aged 15-24 (%)	0	13.787	5.884	2.421
Unemployed (%)	0	30.303	5.771	3.345
Receiving government assistance (%)	1.1	67.9	9.325	6.042
Low income designation (%)	4.213	78.306	17.985	8.175
Median income (thousands, CAD)	11.504	68.736	33.742	9.701
Subsidized housing (%)	0	90.9	7.864	16.833
Houses under major repair (%)	0	41.159	6.423	4.739
Rented households (%)	0	100	47.669	22.77
Post-secondary education (%)	15.337	87.924	53.759	14.423

n = 978

Table 3. Bivariate correlations for independent variables

	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀
Aboriginal (%), X ₁	1	0.06	-.284**	-.301**	-.096**	.113**	.358**	0.05	-.260**	-0.005
Ethnic heterogeneity, X ₂		1	-.195**	-.237**	-.156**	-.165**	.104**	0.048	-0.033	0.038
Visible minorities (%), X ₃			1	.906**	.281**	-.232**	-.322**	.471**	.605**	.097**
Immigrants (%), X ₄				1	.387**	-.224**	-.296**	.429**	.553**	.115**
Recent immigrants (%), X ₅					1	.253**	0.028	0.01	.203**	0.042
Moved within 1 year (%), X ₆						1	.324**	-.283**	-.188**	0.007
Single persons (%), X ₇							1	0.017	-.362**	0.054
Lone parents (%), X ₈								1	.359**	.137**
Males aged 15-24 (%), X ₉									1	.109**
Unemployed (%), X ₁₀										1
Receiving government assistance (%), X ₁₁										
Low income designation (%), X ₁₂										
Median income (thousands, CAD), X ₁₃										
Subsidized housing (%), X ₁₄										
Houses under major repair (%), X ₁₅										
Rented households (%), X ₁₆										
Post-secondary education (%), X ₁₇										

* $p < 0.05$, ** $p < 0.01$

Table 3. Bivariate correlations for independent variables, continued

	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	X ₁₇
Aboriginal (%), X ₁	.131**	.101**	-0.028	.323**	.266**	.381**	0.048
Ethnic heterogeneity, X ₂	-0.041	.109**	0.021	.201**	-0.011	.063*	-.186**
Visible minorities (%), X ₃	.500**	.102**	-.663**	-.231**	-.251**	-.400**	-.602**
Immigrants (%), X ₄	.467**	.182**	-.600**	-.168**	-.238**	-.372**	-.479**
Recent immigrants (%), X ₅	0.02	.279**	-.149**	-.099**	-.095**	.110**	0.055
Moved within 1 year (%), X ₆	-.311**	.208**	.203**	0.004	0.031	.404**	.497**
Single persons (%), X ₇	.186**	.420**	-.152**	.409**	.227**	.733**	.188**
Lone parents (%), X ₈	.560**	.188**	-.557**	.172**	0.006	-.103**	-.541**
Males aged 15-24 (%), X ₉	.180**	0.047	-.423**	-.289**	-.204**	-.424**	-.445**
Unemployed (%), X ₁₀	.142**	.230**	-.213**	.064*	-0.023	0.035	-.113**
Receiving government assistance (%), X ₁₁	1	.197**	-.834**	.251**	0.059	.102**	-.705**
Low income designation (%), X ₁₂		1	-.391**	.411**	0.03	.386**	0.028
Median income (thousands, CAD), X ₁₃			1	-.118**	0.047	-0.048	.725**
Subsidized housing (%), X ₁₄				1	.224**	.450**	0.023
Houses under major repair (%), X ₁₅					1	.259**	.106**
Rented households (%), X ₁₆						1	.269**
Post-secondary education (%), X ₁₇							1

* $p < 0.05$, ** $p < 0.01$

Table 4. Spatial regression results for mischief

	Full residential model	Final residential model	Full ambient model	Final ambient model
Aboriginal (%)	0.003		0.719***	0.87***
Ethnic heterogeneity	0.003		-0.048	
Visible minorities (%)	-0.01		0.002	
Immigrants (%)	-0.042		-0.17**	
Recent immigrants (%)	0.016		0.191	
Moved within 1 year (%)	0.071		-0.184***	-0.187**
Single persons (%)	0.265***	0.28***	-0.068	
Lone parents (%)	-0.085		-0.224**	-0.245**
Males aged 15-24 (%)	0.258		-0.049	
Unemployed (%)	-0.036		-0.037	
Receiving government assistance (%)	-0.123	-0.155**	0.401**	0.37**
Low income designation (%)	0.065		0.23**	0.165*
Median income (thousands, CAD)	-0.029		-0.094	
Subsidized housing (%)	0.02		0.015	
Houses under major repair (%)	0.044		0.028	
Rented households (%)	-0.051**	-0.028*	-0.038	
Post-secondary education (%)	-0.066	-0.067*	0.141**	0.151***
Pseudo R^2	0.217	0.174	0.116	0.105

n = 978, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Spatial regression results for theft from vehicle

	Full residential model	Final residential model	Full ambient model	Final ambient model
Aboriginal (%)	-0.127		0.401	
Ethnic heterogeneity	-0.048		-0.154	
Visible minorities (%)	-0.164***	-0.15***	-0.175	
Immigrants (%)	0.029		-0.123	
Recent immigrants (%)	0.083		0.136	
Moved within 1 year (%)	0.068		-0.239	
Single persons (%)	0.151		-0.532***	-0.406**
Lone parents (%)	-0.181*	-0.238**	-0.48**	-0.374*
Males aged 15-24 (%)	-0.3		-0.695	
Unemployed (%)	-0.043		0.199	
Receiving government assistance (%)	-0.264*		0.603*	1.166***
Low income designation (%)	-0.004		0.249	
Median income (thousands, CAD)	-0.206**	-0.229**	-0.134	
Subsidized housing (%)	0.076*		0.047	
Houses under major repair (%)	0.128	0.139*	0.077	
Rented households (%)	-0.021		0.054	
Post-secondary education (%)	-0.089		0.258*	0.385***
Pseudo R^2	0.168	0.145	0.075	0.054

n = 978, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6. Spatial regression results for theft of vehicle

	Full residential model	Final residential model	Full ambient model	Final ambient model
Aboriginal (%)	0.05		0.249**	0.261***
Ethnic heterogeneity	-0.008	-0.012**	-0.026	
Visible minorities (%)	0.002		-0.002	
Immigrants (%)	-0.024*	-0.024***	-0.042	
Recent immigrants (%)	0.014		0.008	
Moved within 1 year (%)	0.007		-0.049	
Single persons (%)	0.031**	0.042***	-0.026	
Lone parents (%)	-0.01		-0.094*	-0.115**
Males aged 15-24 (%)	-0.057		0.006	
Unemployed (%)	-0.033		-0.035	
Receiving government assistance (%)	0.002		0.181**	0.148**
Low income designation (%)	-0.01		-0.035	
Median income (thousands, CAD)	-0.018		-0.034	
Subsidized housing (%)	-0.001		0.01	
Houses under major repair (%)	0.026	0.031*	-0.008	
Rented households (%)	0.002		0.006	
Post-secondary education (%)	-0.025**	-0.026***	0.026	
Pseudo R^2	0.105	0.091	0.057	0.045

n = 978, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7. Spatial regression results for theft of bicycle

	Full residential model	Final residential model	Full ambient model	Final ambient model
Aboriginal (%)	0.048		0.149	
Ethnic heterogeneity	-0.013		-0.044*	-0.052**
Visible minorities (%)	0.016		-0.059*	-0.099***
Immigrants (%)	-0.084**	-0.07***	-0.07	
Recent immigrants (%)	-0.032		0.02	
Moved within 1 year (%)	0.058		-0.027	
Single persons (%)	0.156***	0.152***	-0.085	
Lone parents (%)	-0.062		0.02	
Males aged 15-24 (%)	-0.133		-0.664**	-0.593***
Unemployed (%)	0.03		0.102	
Receiving government assistance (%)	-0.203***	-0.217***	-0.075	
Low income designation (%)	0.108*	0.115**	0.156**	0.179***
Median income (thousands, CAD)	0.003		-0.08	
Subsidized housing (%)	0.031		0.008	
Houses under major repair (%)	-0.029		-0.099	
Rented households (%)	-0.018		0.016	
Post-secondary education (%)	-0.008		0.062*	
Pseudo R^2	0.196	0.173	0.121	0.11

n = 978, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8. Spatial regression results for other theft

	Full residential model	Final residential model	Full ambient model	Final ambient model
Aboriginal (%)	0.317		0.263	
Ethnic heterogeneity	0.138*	0.118*	0.043	
Visible minorities (%)	0.056		0.006	
Immigrants (%)	0.145	0.202*	0.066	
Recent immigrants (%)	0.347		0.227*	0.296**
Moved within 1 year (%)	0.153		-0.053	
Single persons (%)	0.61***	0.671***	0.13*	0.188***
Lone parents (%)	-0.106		0.112	0.186**
Males aged 15-24 (%)	0.552		-0.435**	-0.549**
Unemployed (%)	-0.663*		-0.017	
Receiving government assistance (%)	-0.267	-0.527**	-0.264**	-0.247***
Low income designation (%)	0.275		0.082	
Median income (thousands, CAD)	0.322*		0.107	
Subsidized housing (%)	-0.047		-0.019	
Houses under major repair (%)	0.075		0.014	
Rented households (%)	0.026		0.045	
Post-secondary education (%)	-0.048		-0.007	
Pseudo R^2	0.051	0.041	0.047	0.04

n = 978, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$