Catalogue no. 36-28-0001
ISSN 2563-8955

Economic and Social Reports

Exploring property crime and business locations: Using spatial analysis and firm count data to reveal correlations in Toronto, Ontario

by Matthew Brown, Mark Brown and Ryan Macdonald

Release date: November 27, 2024

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

• Statistical Information Service	1-800-263-1136
• National telecommunications device for the hearing impaired	1-800-363-7629
• Fax line	1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

© His Majesty the King in Right of Canada, as represented by the Minister of Industry, 2024

Use of this publication is governed by the Statistics Canada Open Licence Agreement.

An HTML version is also available.

Cette publication est aussi disponible en français.

Exploring property crime and business locations: Using spatial analysis and firm count data to reveal correlations in Toronto, Ontario

by Matthew Brown, Mark Brown and Ryan Macdonald

DOI: https://doi.org/10.25318/36280001202401100001-eng

Abstract

This article presents an exploratory analysis of the relationship between the population, firm counts and average property crime from 2017 to 2020 across the Toronto census metropolitan area (CMA). It combines datasets from different domains—crime, business counts and population data—using 500 m by 500 m spatial grids to explore their relationships. At this scale, residential and business land use can be at least partially separated, allowing the independent association between residential populations, business counts and crime to be measured and mapped across the Toronto CMA. This analysis provides a picture of the spatial pattern of crimes across the CMA, explores and validate the data by establishing expected baseline relationships, and points towards areas for more in-depth analysis to determine the relationship between crime and business outcomes. After accounting for the population of grid squares, a positive association between business counts and crime was found, consistent with previous work. Furthermore, after considering population and firm counts, statistically significant spatial clusters of high (and low) crime rates were found. This work therefore sets the foundation for future analysis that would examine how variations in crime rates across space and time affect business outcomes (e.g., firm profitability and exit).

Keywords: property crime, firms, businesses, spatial crime patterns, geospatial analysis, crime hotspots

Authors

Matthew Brown, Mark Brown and Ryan Macdonald are with the Economic Analysis Division, Analytical Studies and Modelling Branch, Statistics Canada.

Acknowledgment

The authors thank their collaborators at Statistics Canada’s Canadian Centre for Justice and Community Safety Statistics for their essential contributions, as well as the reviewers for their insightful feedback, which greatly enhanced this work. In particular, they would like to thank Simon Baldwin, Mathieu Charron, Samuel Perreault and Stephen Tapp for their helpful comments.

Introduction

Increasingly, data and information across various domains (e.g., social, economic and environmental) are being combined to better understand the relationships between different aspects of society and the economy. Geography provides a natural framework for combining often disparate data that may otherwise have no other linkable characteristics and can reveal patterns that point to underlying socioeconomic processes. To this end, this paper examines the spatial correlations between the location of property crime, firms and the population for the Toronto census metropolitan area (CMA).

The focus on the association between the location of firms and crime is motivated, in part, by the growing body of work pointing to a negative relationship between firm outcomes and crime. Evidence suggests that consumers consider crime when deciding whether to visit a business (Fe & Sanfelice, 2022), that business investment is negatively affected by increasing crime (Acolin et al., 2022; Barbieri & Rizzo, 2023), and that higher levels of violent and property crime in neighbourhoods are associated with higher rates of business failure and mobility (moving away) (Hipp et al., 2019). Conversely, declining property crime is associated with higher neighbourhood-level economic activity (Stacy, Ho & Pendall, 2017). While these findings are not universal (see, for example, Bates & Robb, 2008), the weight of the evidence points towards the negative influence of crime on firm outcomes.

The objective of this paper is not to associate crime with firm outcomes per se. Rather, it takes a step back and gathers evidence on the correlation between the presence of firms and crime at the neighbourhood scale. Specifically, it explores how the presence of firms overlaps with property crime at the neighbourhood level. For large geographic units, such as cities or CMAs, population size may be a sufficient metric for measuring property crime rates. However, within Canadian CMAs, crime rates are not uniformly distributed (Savoie, 2008).[1] At smaller geographic levels, such as local neighbourhoods, this metric becomes especially limited, as crime does not strictly follow population size. Crime also occurs where people work and shop (i.e., in locations where firms operate), adding an additional level of complexity to the measurement of neighbourhood-level crime.[2]

This analysis, therefore, combines reported property crime counts with firm counts from Statistics Canada's business microdata and population counts from the Census of Population. It explores the underlying spatial characteristics of these data geocoded to 500 m by 500 m grid squares—a standard areal unit that can be used to unify different types of data. In doing so, the analysis reveals correlations between the variables using traditional non-spatial correlative analysis techniques, such as linear regression, and spatial bivariate mapping and cluster analysis techniques.

The results demonstrate that reported property crime is positively associated with population levels and firm counts across the grid squares. Additionally, bivariate maps and regression analysis illustrate that including firm counts explains the variation in crime locations across grid squares in ways that the population alone does not. After accounting for the population, the exploratory regression model shows

1. Previous work has found that the type of commercial area is related to the level of crime committed. For example, big-box stores and industrial parks tend to be targeted less by all age groups, compared with other types of commercial areas (Charron, 2009, 2011). Of course, multiple neighbourhood-level factors may condition these effects. For example, the presence of local institutions such as schools can influence the level of property crime in an area, the degree of the influence varying depending on the type of school (Willits, Broidy, & Denman, 2013).
2. More firms in a neighbourhood provide greater opportunity for crime, but more "eyes on the street" may reduce it (see Rosenthal & Urrego, 2023). Although the direction of the relationship between crime and the number of businesses is unclear, simple correlative models tend to show a positive association between crime and economic activity (see Stacy, Ho, & Pendall, 2017).

a statistically significant positive association between property crime and consumer-facing firms (e.g., retail stores). The analysis also identifies the presence of statistically significant spatial clusters of neighbourhoods with high crime (e.g., downtown Toronto), where property crime levels are higher than what would be expected given population size and the number of firms. This confirms that there are non-random processes driving spatial patterns in property crime rates and provides additional motivation for future work to better understand the causes of high crime clusters, particularly in relation to the number of firms within a neighbourhood.

The remainder of the study is structured as follows. Section 2 discusses the data sources and the pre-treatment of the data to produce grid-square-based values that are suitable for analysis. Section 3 describes the basic geospatial patterns found in the property crime, firm-level and population data. Section 4 examines the correlation between property crime, firm and population counts using bivariate maps and measures of spatial clustering (i.e., Local Moran's I) derived from regression residuals. Section 5 concludes the paper.

Data

The analysis takes advantage of three types of data: crime, firm and population counts. This section describes the characteristics and sources of these data and how they are combined geographically through a uniform grid.

Key to the analysis, of course, are measures of crime—specifically, property crime. Property crime is the focus because it is more likely to be associated with businesses than other types of crime, such as homicide or drug trafficking. The property crime dataset used here includes all types of property crime violations under the Canadian *Criminal Code*, including breaking and entering, various forms of theft, possession and trafficking of stolen property, and criminal mischief.[3] This dataset was obtained from Statistics Canada's Canadian Centre for Justice and Community Safety Statistics and includes the geographic point locations of reported crimes at various levels of geography, with the majority being captured at the block-face and dwelling levels.

The crime data were filtered to contain only property crimes and to include only point locations geocoded at finer-scale geographies (i.e., dissemination area level and below). Additionally, geocoded locations of fraud and other virtual crimes often differ from their actual location. The victim's residence is often used as the location, even though it may not always be appropriate, such as in the case of online fraud. Therefore, the following crimes were removed from the analysis: fraud; identity fraud; identity theft; and altering, removing or destroying vehicle identification numbers.

Firm-level data were derived from the Longitudinal Business Database (LBD), which was built using the Business Register—a dataset covering the universe of firms in Canada (Statistics Canada, 2024). The LBD is used to construct the firm count variable, and the location of the enterprise is used to construct the firm count by location variable. Because most firms have only one operating location, this variable reasonably estimates the number of firms in a grid square. A key aspect of the LBD is that it allows researchers to consistently track firms over time, facilitating future work on the relationship between crime and firm outcomes.

Firm counts and crime counts were originally formatted as a spatial points layer and then aggregated into a tessellation of 500 m by 500 m grid squares covering the Toronto CMA. The grid squares present

3. For more details, refer to Statistics Canada (2022).

longitudinal units whose values were averaged over the period from 2017 to 2020 to create a surface of grid squares containing the average annual number of firms and the average number of property crimes for each square. To facilitate the regression analysis, average firm counts were additionally split into, and calculated for, consumer-facing and non-consumer-facing firms. Consumer-facing firms are those with customers as clientele (e.g., retail), as opposed to other businesses. Firms were separated into these categories using their North American Industry Classification System (NAICS) codes, following a classification scheme identified by Kane, Hipp and Kim (2017). The expectation is that consumer-facing firms would be more likely to affect, and be affected by, property crime (e.g., shoplifting).

Population data for the study were sourced from the 2021 Census of Population, made available via GeoSuite (Statistics Canada, 2021c), as well as 2021 geographic boundary files (Statistics Canada, 2021b). Population data at the dissemination block (DB) level were converted into a 500 m by 500 m grid square surface via geometric intersection.[4]

Because of reporting and collection methods, certain grid square locations may have missing data in specific years. To maximize data inclusion, null values were treated as 0 if the average of any variable (i.e., property crime, firms or population) over the period was above 0. Locations with null values for all three variables were excluded from the analysis. For example, a grid square located on an airport runway would be removed from the analysis, but a grid square located in a residential area with at least one resident and no recorded firms or crimes would remain. For brevity, in the remainder of this paper, the term “crimes” refers to the average counts of reported property crime, while the term “firms” refers to the average number of firms, with averages calculated across the four-year study period. Descriptive statistics for each variable are presented in Appendix A. All variables tend to have right-skewed distributions because of, in part, the presence of null values in the data.

Analysis based on spatial characteristics

The analysis focuses on the population of individuals, the population of firms and the reporting of crimes within the boundary of the Toronto CMA. The Toronto CMA is the most populous CMA in Canada, with a population of 6,022,225 and a land area of 5 903 km^2, according to the 2021 Census of Population (Statistics Canada, 2021a).

Within the Toronto CMA, theft under $5,000 (not including motor vehicles) was the most prevalent type of property crime for each year from 2017 to 2021. The next largest categories were fraud,[5] mischief, breaking and entering, and theft of a motor vehicle (Table 1). These five categories constitute most property crimes in the Toronto CMA and represent violations that can directly affect businesses and people.

4. The ratio of the intersected DB to its original area was calculated and then multiplied against the original DB population value. This intersected population value was then summed for all polygons inside the grid square and rounded to the nearest integer to generate a single population value for each grid square. This assumes that the population is evenly distributed throughout the DB. However, this is not necessarily the case, especially in rural areas, where DBs have larger areas by design. Therefore, allocating the data into fixed grid squares introduces some location error into the analysis.
5. As mentioned, virtual crimes such as fraud were removed from the analysis because of concerns with geolocation accuracy.

Table 1
Incident-based crime statistics, by detailed violation, Toronto, Ontario

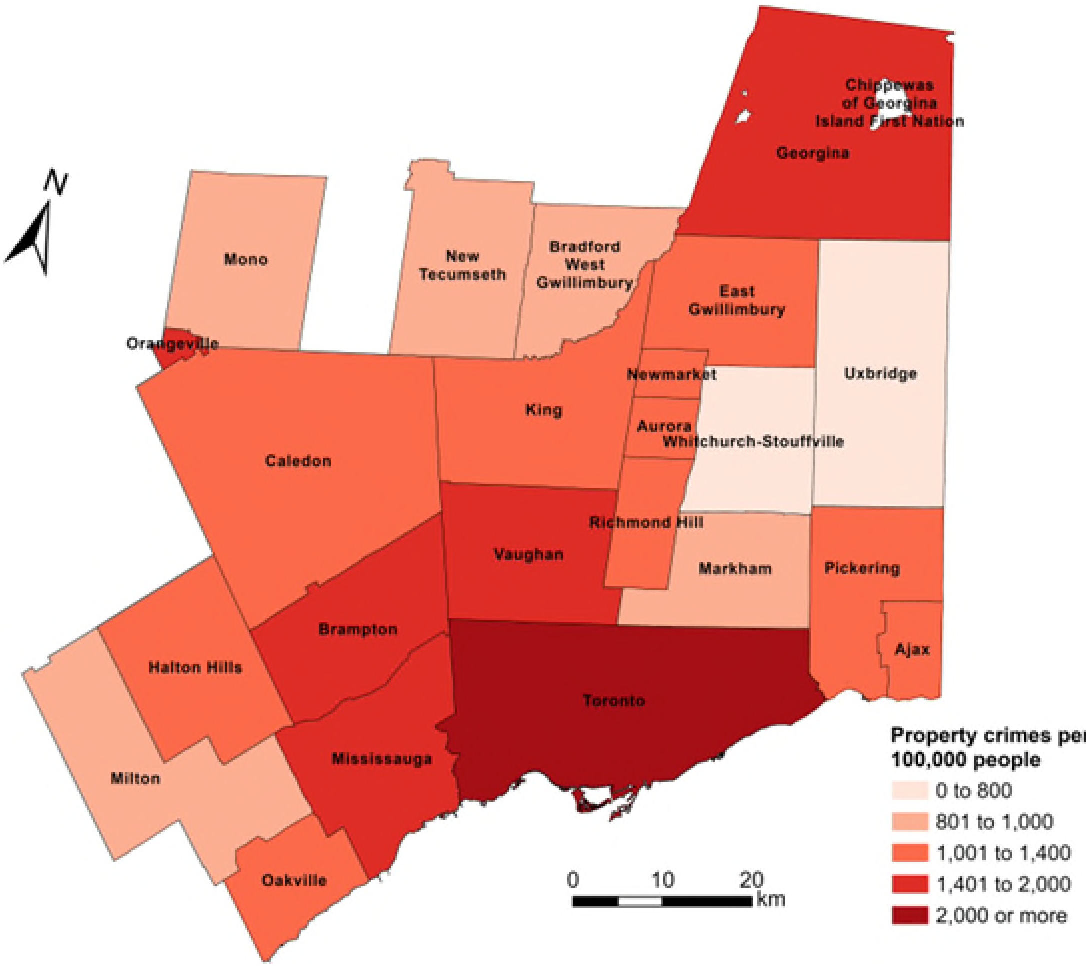
	Number of actual incidents											
Year	Breaking and entering	Possession of stolen property	Trafficking in stolen property	Theft of a motor vehicle	Theft over $5,000 (non-motor vehicle)	Theft under $5,000 (non-motor vehicle)	Fraud	Identity theft	Identity fraud	Mischief	Arson	Altering, removing or destroying a vehicle identification number
2017	13,493	1,313	49	8,014	2,300	67,009	15,892	117	2,227	16,459	418	8
2018	14,300	1,560	58	9,971	2,500	77,075	18,395	100	2,123	16,405	372	3
2019	14,981	1,451	42	10,641	2,501	76,928	21,614	159	2,042	15,977	319	0
2020	11,614	1,578	69	11,509	2,231	57,794	19,435	145	2,189	16,219	408	2
2021	9,748	1,347	95	14,021	2,277	59,868	18,229	87	2,390	16,150	383	4

Notes: This table adjusts population values w ithin census metropolitan areas (CMAs) to reflect actual policing boundaries, meaning the values show n in this table are not necessarily reflective of the Statistics Canada population for the Toronto CMA as a w hole. The follow ing crimes w ere excluded from the analysis: fraud; identity theft; identity fraud; and altering, removing or destroying a vehicle identification number.

Source: Statistics Canada, table 35-10-0177-01.

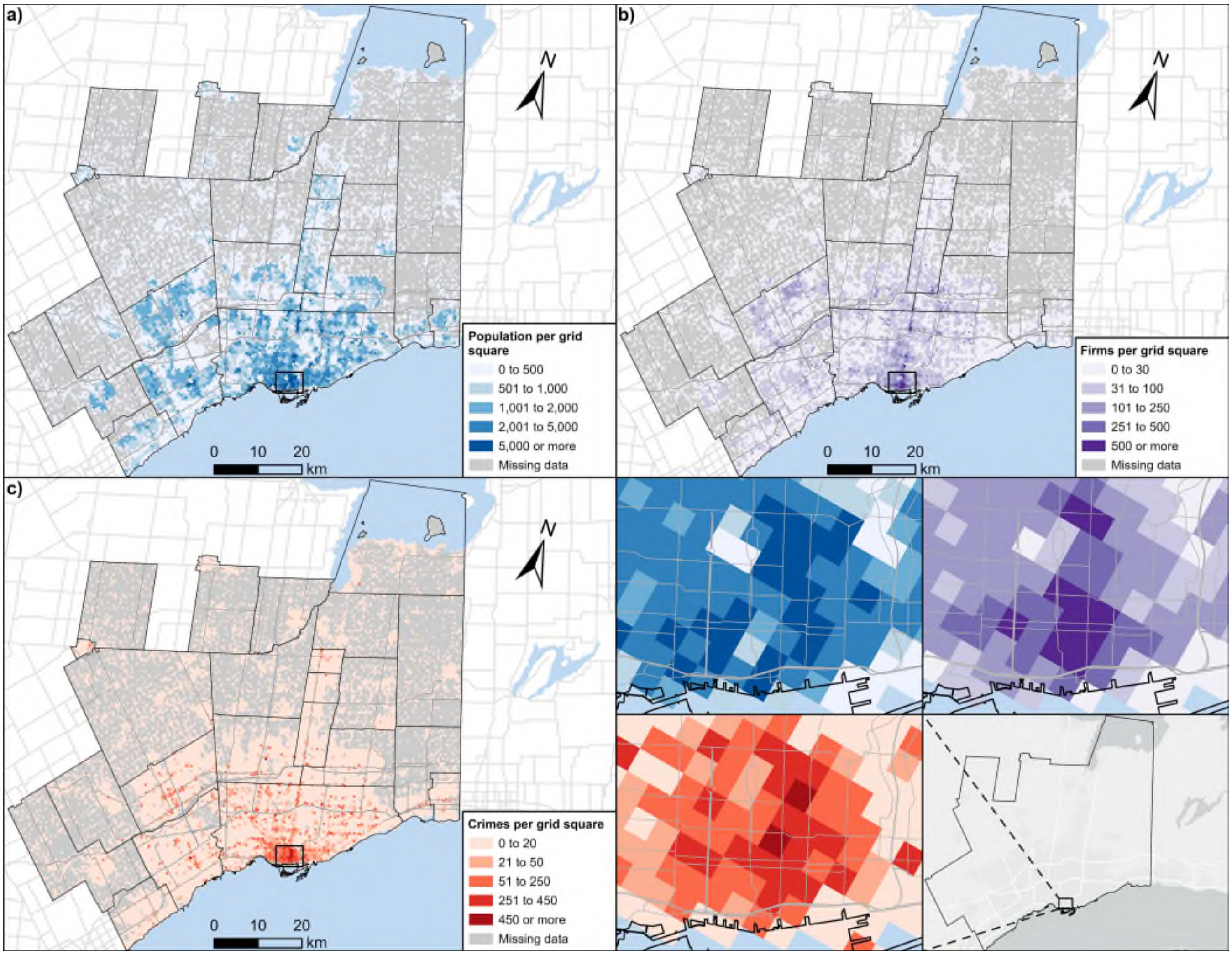
To examine reported crimes at more disaggregated, regionally specific geographies, data are often reported rates, such as crime occurrences per 100,000 people, and over areas, such as counties, provinces or sections of a city. Map 1 provides an example of this type of crime rate based on the average annual level of crime from 2017 to 2020 across census subdivisions (CSDs) within the Toronto CMA. While this presents a coarse look at the dispersion of reported property crimes, insights can be drawn from this exercise. The highest property crime rates are in the Toronto CSD, with lower but highly variable rates in the surrounding areas. This scale, however, likely masks considerable variation in crime rates within CSDs. Residents often experience crime at the neighbourhood level. This is also true for firms. Firms tend to concentrate in districts (e.g., because of zoning or various locational advantages common across firms). Therefore, to examine the relationship between firms, the population and crime rates, a scale of analysis that captures this spatial variation is required. To address this, the spatial distribution of property crime is reported at a finer spatial scale, using standardized grid squares as a linking geography. Maps showing the distribution of the population, firms and crimes over the study area in grid squares are shown in Map 2.

Map 1
Average property crime rate by census subdivision, Toronto, Ontario, 2017 to 2020



Notes: Virtual crimes have been removed via filtering from the property crime statistics. All census subdivisions in Toronto are included, except for Chippewas of Georgina Island First Nation because of data limitations.
Source: Statistics Canada, authors' calculations.

Map 2
Overview maps showing a) the population, b) average firm counts and c) average property crime counts inside grid squares across the Toronto census metropolitan area



Note: Grid squares are labelled as "missing data" if they have null values in all three variables across the four-year period.
Source: Statistics Canada, authors' calculations.

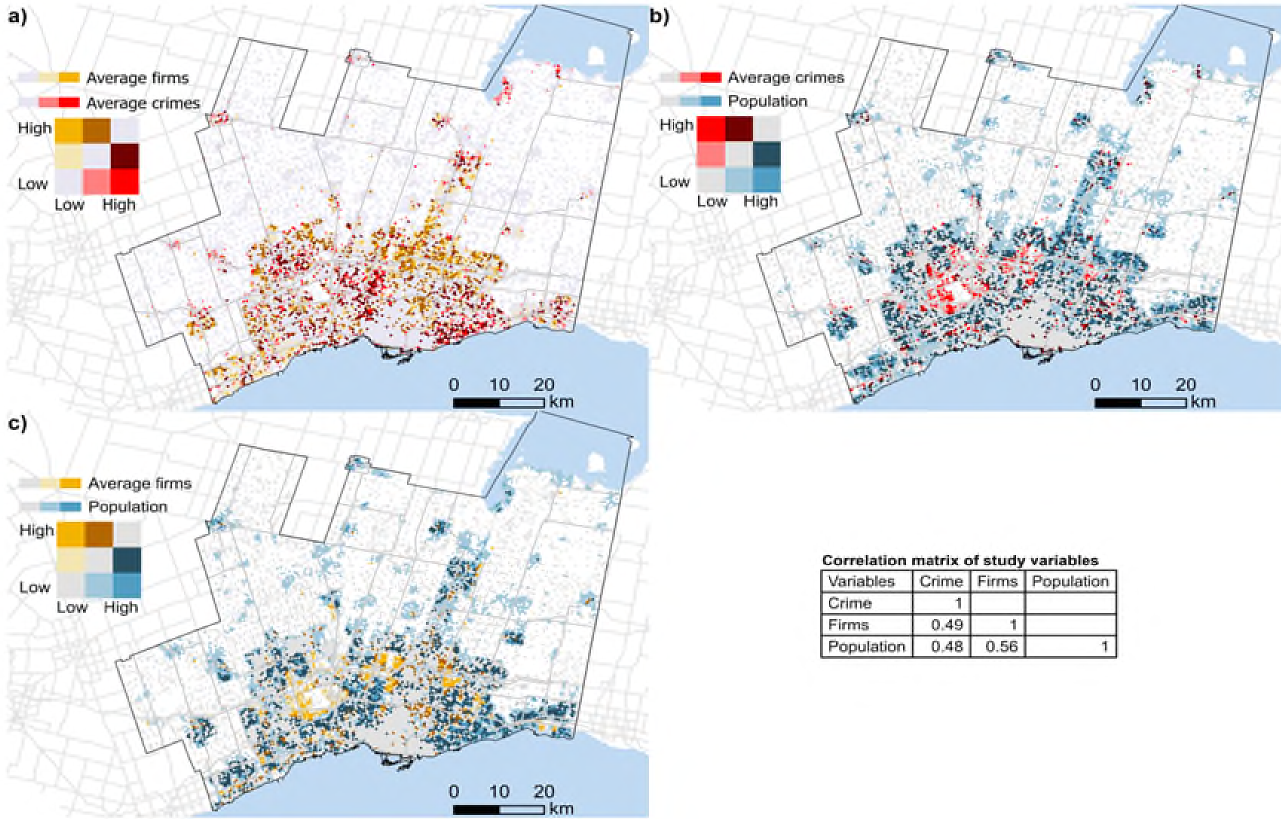
All three maps in Map 2 illustrate a common concentration of population, firms and crime in downtown Toronto (see also the inset maps), but there are different geographic patterns in their detailed geographies. As illustrated in Map a), the population is concentrated in the Toronto CSD, peaking in the downtown core. In Map b), the average number of firms is also highest in the Toronto CSD, with its highest levels in the downtown core at Union Station and extending northward along Yonge Street. High concentrations of firms can also be observed in downtown Mississauga; Brampton; and near major highway intersections in Vaughan, Richmond Hill and Markham. In Map c), a high concentration of crime occurs in the central downtown core. Upon closer inspection, squares with higher average property crime tend to follow the road network and are generally in CSDs immediately surrounding Toronto to the north and west. Areas of high crime appear to be much more locally concentrated, compared with what is observed in the population or firm maps. For example, high-crime areas in Brampton and Mississauga are in their downtown cores and do not appear to spread out as extensively into surrounding areas. This is consistent with previous research showing that, for instance, shoplifting was not found to be affected by the characteristics of adjacent neighbourhoods in Toronto (Charron, 2009). Looking at the univariate choropleth maps in Map 2, there is an apparent high degree of association between the locations of reported property crimes and population counts and firm locations. While this observation is intuitive, the individual maps do not allow for a statistical analysis of how crime locations, people and firms interact. Doing so requires multivariate analysis techniques.

Correlation analysis

As the informal comparison of the maps in Map 2 suggests, when the counts of crime, firms and population across grid squares are compared, there is a relatively strong positive correlation among them. The Pearson correlation coefficient (r) between crime and firms is 0.49. Likewise, the r between property crime and the population is 0.48, while between firms and the population it is 0.56 (see Map 3). While these coefficients indicate that the data are correlated and reinforce the observations in the univariate maps, they conceal local variations in correlation that become apparent in bivariate maps. Of particular interest are locations where there is disagreement, such as areas where crime is relatively high, but the population is relatively low, and whether these areas visually correspond to locations where firms are more prevalent.

Bivariate maps provide a visual representation of the correlation between two variables. In these maps, each variable is divided into three bins, potentially creating nine unique colour classes. The low to high values of each variable can be read from bottom to top (Variable 1) and left to right (Variable 2). The upward diagonal from bottom left to top right represents a positive correlation in the data, whereas the off-diagonal colour ramps indicate increasing disagreement between the variables. Because a positive correlation among the variables was established, the left-to-right diagonal is identified by a light-grey colour to highlight the off-diagonal cells, which are areas where this relationship does not hold. For example, the colour ramp from the bottom left to the top left (i.e., from light grey to light yellow to dark yellow, or from light grey to light red to dark red) represents areas that show increasing values in Variable 1, while remaining low in Variable 2. The opposite disagreement pattern (i.e., where the values in Variable 2 become increasingly high, while remaining low in Variable 1) can likewise be observed along the bottom row, from left to right. Map 3 displays the bivariate maps of every possible variable interaction, along with the corresponding correlation coefficients between variables.

Map 3
Bivariate maps comparing a) firms and crimes, b) the population and crimes, and c) the population and firms



Correlation matrix of study variables

Variables	Crime	Firms	Population
Crime	1		
Firms	0.49	1	
Population	0.48	0.56	1

Source: Statistics Canada, authors' calculations.

In terms of patterns that emerge from the bivariate maps, areas where crime, firms and the population are positively correlated are seen in generally the same locations across each map (e.g., downtown Toronto). Of particular interest are the patterns of variable disagreement represented by the colours on the off diagonals. Specifically, areas where the average number of firms and the average number of crimes appear to be high, while the population is low—i.e., red squares in Panel b) and yellow squares in Panel c)—appear in the areas surrounding Toronto/Lester B. Pearson International Airport, MacMillan Yard and various commercial locations (e.g., Etobicoke City Centre). This variable disagreement reveals that the presence of many firms does not necessarily correspond to a high population and that low-population areas can still be locations where crime occurs. For example, for the grid squares that surround Toronto/Lester B. Pearson International Airport, there is little relationship between the population and crimes, as shown by the red squares in Panel b), but there is a positive relationship between firm counts and crimes.

Regression modelling

At root, the bivariate choropleth maps demonstrate that crime and firm or population counts are positively related, but that firm counts and the population potentially have independent associations with crime. To formally test this, the following regression model is estimated using ordinary least squares (OLS):

$$Crime_i =^o \beta_o +^o \beta_1 CFfirms_i + \beta_2 NCFfirms_i + \beta_3 population_i +^o \varepsilon_i$$

Table 2
Ordinary least squares regression model estimates of the count of property crimes as a function of firm and population counts across grid squares

Dependent variable: Average number of property crimes	Coefficient	*t*-statistic	
(Intercept)	-2.628	-3.239	**
Average count of consumer-facing firms	1.010	8.796	***
Average count of non-consumer-facing firms	-0.040	-1.337	
Population count	0.011	5.434	***
Observations	...	12,130	
Adjusted R²	...	0.35	
Residual standard error	...	29.61 (df = 12,160)	
F-statistic	...	3,171 (df = 12,160)	

** significantly different from zero ($p < 0.01$)

*** significantly different from zero ($p < 0.001$)

Notes: df = degrees of freedom. Robust standard errors (HC3, or heteroskedasticity-consistent standard error estimator, version 3) were used for t-statistics and model significance.

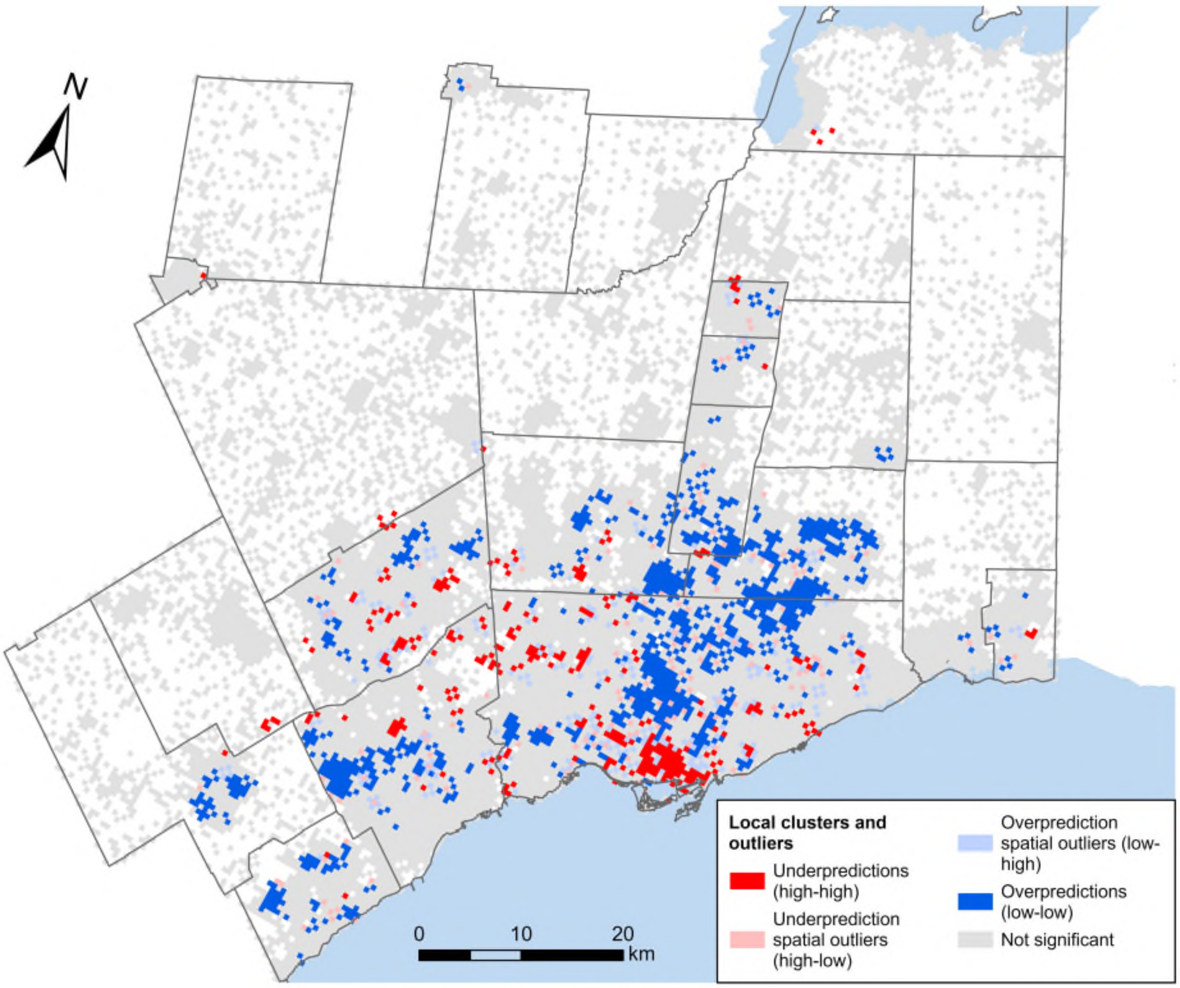
Source: Statistics Canada, authors' calculations.

In line with the existing literature, the model reveals a statistically significant positive relationship between consumer-facing firms and property crime, after taking the population into account. This finding supports the notion that these firms are potentially more exposed to crime because they serve the broader public, putting them at greater risk (e.g., from shoplifting). This is further reinforced by the statistically insignificant association between counts of non-consumer-facing businesses and crime. The population has the expected positive independent association with crime. Together, the independent variables account for 35% of the variation in the average number of property crimes. This is reasonably high given how few variables are included in the model.

Beyond the independent associations between population and firm counts and crime, the residuals of the model are also of interest. They measure the degree to which crime is higher or lower than would be expected after accounting for the count of people and firms in a grid square. If there are spatial clusters of grid squares where crime is higher (or lower) than expected, then this suggests there are underlying spatially non-random factors that result in higher (or lower) crime that may warrant further investigation. For instance, clusters of positive residuals—where crime is higher than expected—may result if there are clusters of grid squares with firms that are particularly vulnerable to property crime (e.g., retail stores). However, if there are no statistically significant spatial patterns in the residuals, the null hypothesis—that the factors leading to higher or lower levels of crime than expected are randomly distributed across space—cannot be rejected.

As might be reasonably expected for such a simple model, the residuals are not spatially random. The Global Moran's I, which ranges from -1 to +1, yielded a value of 0.20 (z-score = 34.44), suggesting strong positive spatial autocorrelation and rejection of the null hypothesis of spatial randomness.[6] Moreover, the Local Moran's I based on the model residuals indicates statistical clusters in the study area. The resulting map of clusters is presented in Map 4.

Map 4
Local Moran's I cluster map of ordinary least squares model residuals



Source: Statistics Canada, authors' calculations.

From this map, model underpredictions of crime (i.e., groupings of relatively high-value positive residuals, shown in red) cluster in downtown Toronto in areas with the highest density of firms or people. They also appear sparsely around the CMA and particularly within the Toronto, Etobicoke and Brampton CSDs, primarily along major roads and highways. In general, clusters of underpredictions tend to appear in areas used heavily for commercial purposes, such as shopping centres or downtown commercial districts. Conversely, several large clusters of model overpredictions (i.e., groupings of low-value

6. Moran's I is a measure of spatial association. In its global form, a statistically significant value above 0 indicates that grid squares with above-average crime values tend to be located near other grid squares with above-average values (four closest neighbours), while those with below-average values tend to be located near other grid squares with low values. The Local Moran's I tests whether individual grid squares are near others with similarly above-average values (high-high) or below-average values (low-low) (see Map 4). In this sense, it provides a test of spatial clustering that can be mapped.

residuals, shown in dark blue) can be found in many neighbourhoods across the CMA, but particularly in North York, Mississauga, Markham, Vaughan and Richmond Hill. These overpredictions typically occur in regions of heavy residential land use. This finding suggests the presence of other underlying spatially non-random factors that result in higher (or lower) crime counts.

Lastly, the map also identifies statistically significant spatial outliers. Underprediction spatial outliers (shown in pink) are grid squares with above-average residual values whose neighbours, on average, have lower values than would be expected under spatial randomness (e.g., a high-value residual grid square surrounded by low-value residual grid squares). The opposite holds true for overprediction spatial outliers (shown in light blue). Underprediction spatial outliers tend to appear in locations with locally high levels of commercial activity (such as grid squares containing big-box retail locations) that are adjacent to areas of low activity, such as parks, highways or residential areas. Conversely, overprediction spatial outliers are often in residential areas situated on the periphery of commercial areas.

Limitations and assumptions

While this article was intended as an exploratory analysis, the limitations of the data need to be acknowledged. Because the analysis uses averaged data from a four-year panel dataset of crime and firm counts, these counts may be underestimated in grid squares that saw significant growth over the period, such as those on the outskirts of Toronto. Another limitation of the dataset pertains to how data are captured in rural locations. Rural postal codes cover much larger areas than their urban counterparts. As the firm data are based on centroids of postal code geography during the grid square aggregation, grid squares with no actual firms within them could be identified as having firms. Lastly, as the COVID-19 pandemic may have affected crime rates, the analysis was repeated excluding 2020 data. The results were qualitatively unchanged, so the paper includes data from the entire period (2017 to 2020).

As illustrated by the spatial analysis of the OLS regression residuals, several other variables beyond population size and the number of firms likely influence the observed crime counts. Socioeconomic factors or local accessibility probably also play a role. Another consideration is the reliance solely on firm counts in the analysis, without accounting for variations in firm size. Consequently, the model treats small businesses (i.e., firms with 1 to 99 employees) with the same weight as large enterprises (i.e., firms with 500 or more employees), potentially resulting in bias. As most businesses in Canada are small (e.g., approximately 87% of Ontario firms in 2020 had fewer than 20 employees [Statistics Canada, 2021d]), the results tend to reflect the relationship between small firms and property crime.

Conclusion

This paper presents an exploratory analysis of local property crime patterns in the Toronto CMA. Through the aggregation of crime, population and firm count data into a uniform spatial grid dataset, spatial analysis techniques were used to explore patterns in the data. Although property crime, population and firms are shown to be positively correlated with each other, the correlation varies over space. Bivariate maps highlight the spatial correlations between the different variables and, specifically, the utility of including a firm count variable. Moreover, the presence of statistically significant spatial clusters in the data illustrates that there are grid squares exhibiting levels of crime that are either higher or lower than would be expected given the population and number of firms there. Further exploration is warranted to understand what additional variables should be included, as well as to test different types of models (e.g., count-based models).

This paper illustrates that using a grid square geography in conjunction with firm count information reveals an independent association between the presence of firms and crime. To fully understand neighbourhood-level crime, the firm dimension needs to be considered. As the presence of firms is important for understanding neighbourhood crime, neighbourhood-level crime may also be important for understanding firm outcomes (Hipp et al., 2019; Stacy, Ho & Pendall, 2017). Therefore, an avenue for future work would be to explore how crime influences firm outcomes, like profitability and exit rates. This would fully leverage the underlying data that track firms over time.

Appendix A

Summary statistics

Appendix Table A.1

Summary statistics for variables of interest across grid squares

	Average firms	Average crimes	Population (2021)
Minimum	0	0	0
Maximum	680	500	10,969
Mean	25	9	502
Median	10	3	165
Number of observations			12,130

Notes: The minimum and maximum of the average firms variable and the average crimes variable are calculated as the average of the 25 bottom and top grid squares, respectively. For the median, the 12 observations ranked below and above its value and the median itself are averaged. These summary statistics include only grid squares that contained a firm or a crime over the study period.

The mean and median values of each variable are much closer to the minimum value than to the maximum value. This can be explained by the presence of a large number of null values in the dataset, particularly for property crime. All the variables in this analysis have a right-skewed distribution.

References

Acolin, A., Walter, R. J., Skubak Tillyer, M., Lacoe, J., & Bostic, R. (2022). Spatial spillover effects of crime on private investment at nearby micro-places. *Urban Studies, 59*(4), 834-850. https://doi.org/10.1177/00420980211029761

Barbieri, N., & Rizzo, U. (2023). The impact of crime on firm entry. *Journal of Regional Science*, *63*(2), 446-469. https://doi.org/10.1111/jors.12628

Bates, T., & Robb, A. (2008). Crime's impact on the survival prospects of young urban small businesses. *Economic Development Quarterly*, *22*(3), 228-238. https://doi.org/10.1016/j.jeconbus.2007.09.004

Charron, M. (2009). *Neighbourhood characteristics and the distribution of police-reported crime in the city of Toronto*. Statistics Canada. https://www150.statcan.gc.ca/n1/pub/85-561-m/85-561-m2009018-eng.htm

Charron, M. (2011). *Neighbourhood characteristics and the distribution of crime in Toronto: Additional analysis on youth crime*. Statistics Canada. https://www150.statcan.gc.ca/n1/pub/85-561-m/85-561-m2011022-eng.htm

Fe, H., & Sanfelice, V. (2022). How bad is crime for business? Evidence from consumer behavior. *Journal of Urban Economics, 129*, 103448. https://doi.org/10.1016/j.jue.2022.103448

Hipp, J., Williams, S. A., Kim, Y. A., & Kim, J. (2019). Fight or flight? Crime as a driving force in business failure and business mobility. *Social Science Research*, *82*, 164-180. https://doi.org/10.1016/j.ssresearch.2019.04.010

Kane, K., Hipp, J., & Kim, J. (2017). "Analyzing accessibility using parcel data: Is there still an access-space tradeoff in Long Beach, California?" *The Professional Geographer* 69(3), 486-503. https://doi.org/10.1080/00330124.2016.1266951

Rosenthal, S. S., & Urrego, J. A. (2023). Eyes on the street, spatial concentration of retail activity and crime. *Working paper, Syracuse University.*

Savoie, J. (2008). *Analysis of the spatial distribution of crime in Canada: Summary of major trends.* Statistics Canada. https://www150.statcan.gc.ca/n1/pub/85-561-m/85-561-m2008015-eng.htm

Stacy, C. P., Ho, H., & Pendall, R. (2017). Neighborhood-level economic activity and crime. *Journal of Urban Affairs, 39*(2), 225-240. https://doi.org/10.1111/juaf.12314

Statistics Canada. (2021a). *Focus on geography series, 2021 Census of Population*. https://www12.statcan.gc.ca/census-recensement/2021/as-sa/fogs-spg/Page.cfm?lang=e&topic=1&dguid=2021S0503535

Statistics Canada. (2021b). *Boundary Files*. https://www150.statcan.gc.ca/n1/en/catalogue/92-160-X

Statistics Canada. (2021c). *GeoSuite*. https://geosuite.statcan.gc.ca/geosuite/en/index

Statistics Canada. (2021d). Table 33-10-0304-01 *Canadian Business Counts, with employees, December 2020.* https://doi.org/10.25318/3310030401-eng

Statistics Canada. (2022). Table 35-10-0177-01 Incident-based crime statistics, by detailed violations, Canada, provinces, territories, census metropolitan areas and Canadian Forces Military Police. https://doi.org/10.25318/3510017701-eng

Statistics Canada. (2024). *Business Register*. https://www.statcan.gc.ca/en/survey/business/1105

Willits, D., Broidy, L., & Denman, K. (2013). Schools, neighborhood risk factors, and crime. *Crime & Delinquency, 59*(2), 292-315. https://doi.org/10.1177/0011128712470991