# Exploring Socio-Economic and Built-Environment Determinants of Alcohol Expenditure for the City of Toronto: A Spatial Analysis Approach

by

Andrew Sik-On Leung

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Health Studies and Gerontology

Waterloo, Ontario, Canada, 2016

## AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

# Abstract

Alcohol expenditures can provide a different perspective on alcohol use by providing an understanding of the consumer demand for alcohol, the effect of alcohol sales on the economy and the impact on the household budget. Previous studies have focused on alcohol expenditure at the individual-level and have not considered population level factors or the influences of geographical variation. The goal of this study was to examine the socio-economic and built-environment characteristics associated with alcohol expenditures at the small-area level in the City of Toronto.

Alcohol expenditure data consisting of purchases in licensed premises and purchases in stores for the year 2010 were retrieved from the Survey of Household Spending (SHS) at the Dissemination Area (DA) level. Socio-economic and built-environment variables were retrieved from the 2006 Census of Canada and DMTI Enhanced Points of Interest (EPOI) data, respectively. Multivariate spatial regression models were used to analyze the associations between alcohol expenditure and both socio-economic and built-environment variables (i.e. alcohol outlet density, restaurant density and subways).Global Moran's I identified geographic variation for both types of alcohol expenditures. Local Moran's I identified three hot-spots and three cold-spots for licensed premises expenditures whereas four hot-spots and two cold-spots were identified for purchases in stores expenditures. The spatial Durbin error model was identified as the best spatial model for both types of alcohol expenditures. For licensed premises expenditures, positive associations were found with seven socio-economic variables and three built-environment variables. For purchased in stores expenditures, positive associations were found with three socio-economic variables and subway intercepts.

This study was the first to identify socio-economic and built-environment characteristics associated with alcohol expenditures at the DA level for the City of Toronto. Additionally, these findings highlight the importance of examining associations between built-environment variables and health behaviours, as well as the importance of considering the actual geography of the study region. Finally, the spatial regression models, visualizations and Geographical Information System (GIS) methods employed in this study can be collectively applied as a toolkit to study other health behaviours that present with geographical variation.

Future studies should examine additional types of built-environments for associations with alcohol expenditures. To create an even more complete understanding of what drives alcohol expenditures, the findings from this study should be combined with individual-level characteristics in spatially-explicit multilevel models.

## Acknowledgments

My journey through graduate school to complete this thesis has been both exhilarating and frustrating, as well as everything in-between. There were many nights when I was despondent and despaired of whether I would even finish. I certainly would not have been able to complete my greatest endeavour to date without the great support, guidance and belief of my supervisor, my committee members and my family and friends.

**To Dr. Jane Law:** Thank you for taking a chance on me at the beginning and for helping me navigate spatial analysis and GIS from the very beginning. I deeply appreciate all of your mentorship and patience in seeing me through this process. It has been my greatest privilege to learn from you. Thank you!

**To Dr. Martin Cooke:** Thank you firstly for letting me work with you on the mapping project, so I had the good fortune of meeting you! Your expertise and perspective were invaluable in helping me grow. Thank you so much for always having time just to chat with me and for being so encouraging and so helpful!

**To Dr. Scott Leatherdale:** I am so glad I came in to see you about my HSG 601 assignment! Your advice and perspective has really helped to shape my thought process. Thank you so much for all of the advice, feedback and encouragement throughout the entire process!

**To Mom and Dad:** I know it has taken me longer than any of us expected, but without your abundant support and love, I would not have been able to finish this project. You both drive me to succeed and to strive for more. Thank you and love you both always!

**To Jianhua Zhao, Sanjeevan Thayaparan, Matt Wawrykow, Matt Stein, Bryan Choi, Michael Nguyen, Ryan Rozario, Joe Li, Arsalan Afzal and Timothy Toong:** You guys have been everything and more to me. There is no way I would have been able to get through this process without the long talks, the advice you gave me and your constant support. You guys kept pushing me and keeping me honest. Thank you brothers!

**To Matt Quick, Henry Luan and Yujie Yang:** Your guidance with GIS issues and your company as fellow students of Dr. Law's was invaluable. Thank you all.

**To my brother, Philemon Leung:** Thank for the being the best roommate for two and a half years. I could not be more thankful for all of the great times we shared, the desperate moments we had and the McDonald's runs we went on! You have brought the best out of me and have always had my back. It means the world to me to have your unwavering support throughout this thesis writing process. Never forget brother: "War within, War without, War unending!" On to new things we go!

**To Heather Hofstetter:** Thank you so very much for putting up with my mood swings throughout this journey. There were so many times that I felt down and out, and yet, you kept me going. You have been extremely patient with me and have given me all of your support. I am so very grateful to be with you. Love you lots Dear!

"I can do all things through Christ who strengthens me". (Phil 4:13) – Thank you Lord Jesus Christ!

## List of Figures

## List of Tables

# Chapter 1: Introduction

Alcohol use is a significant public health challenge that has been causally linked to over two hundred different types of acute and chronic harm (2, 3). The World Health Organization (WHO) defines harmful alcohol use as a pattern of use that leads to physical and mental injury, in addition to detrimental social consequences (4). Yet, in order for alcohol use to be characterized as harmful, several aspects of drinking need to be considered: 1) Volume of alcohol consumed over time; 2) Frequency of drinking to the point of intoxication; 3) Drinking which occurs in a context that poses public health risks; and 4) The quality of alcohol that is consumed (5).

Harmful alcohol use is also recognized as one of the four main modifiable behavioural risk factors of non-communicable diseases and was responsible for 8, 953 deaths and 172, 255 potential years of life lost in Canada in 2005 (6). In Ontario, indirect and direct costs due to alcohol have been estimated to cost the province $5.3 billion per year (3, 7, 8). Despite the myriad of negative consequences, there are still some positive side effects linked to the consumption of low levels of alcohol – including protective effects against cardiovascular diseases – that complicate the discussion for public health officials (3, 9-12). Taking into account these various considerations, low-risk drinking guidelines have been established in Canada to help balance the potential benefits of low level consumption while mitigating the significant increase in the risk of adverse effects at high levels of alcohol consumption (13, 14). Despite such public health initiatives, many Canadians continue to drink in excess of the recommended low-risk guidelines and are at elevated risk of experiencing alcohol-related harm. This also creates unintended economic and societal costs[1] which are only projected to rise with increasing levels of consumption (3).

Analysis of such costs is aided by the measurement of alcohol use in two primary forms: alcohol consumption and alcohol expenditure. The majority of alcohol use studies to date have focused on alcohol consumption, with multiple individual-level studies examining the associations between differing levels of consumption, various alcohol harms, and individual-level predictors of alcohol use. Alcohol expenditure, while similar to alcohol consumption, has been examined far less frequently and can also provide an opportunity to examine alcohol use from a different but equally important perspective. Using expenditure can provide insight into the economic impact of alcohol use as well as spending patterns of individuals and households (15). Additionally, alcohol expenditure can also be examined in the context of household income and spending (16).

Previous studies in the United States, Canada and Asia have examined alcohol consumption and socio-economic determinants at the individual-level using household survey data, thereby establishing this association as it relates to alcohol consumption at the individual-level (15, 17-19). This focus on the individual-level of analysis however, ignores both the contextual effects of geography such as surrounding areas and environmental features as well as socio-economic factors operating at the area-level. Both geography and area-level socio-economic may be important factors for understanding and modifying alcohol expenditure behaviour at the population level as they have been for other health behaviours (20-22).

[1] Social costs are (negative) consequences are borne by society in addition to effects experienced by the individual.

Often, social phenomena such as risky health behaviours like alcohol consumption have a tendency to aggregate in particular geographic areas (23). This clustering of health behaviours shows that the location of an individual is also a vital component in understanding the association between alcohol use and socio-economic determinants. The environment can also potentially influence alcohol use through the built-environment. The built-environment is the physical environment that is constructed by human activity (24). For alcohol use, examples of pertinent built-environment elements that may possibly influence how people access alcohol include alcohol retail stores, restaurants retailers (outlets), subways as well as commercial and residential land use. Additionally, associations between individual-level factors and health behaviours frequently also often do not hold in the same manner and magnitude at an area level (25, 26). Therefore, studying area level socio-economic factors can reveal novel and important associations with alcohol expenditures.

Small-area level studies can potentially help to account for these geographical and environmental influences; however, these types of studies do not exist in the alcohol expenditure literature. Using a small-area level study, the elements of geography can be incorporated in a quantitative fashion by applying spatial analysis methods. Spatial analysis utilizes data that includes geographical coordinates known as spatial data to study the effects of geography on the variable of interest (i.e. alcohol expenditure) using geographical information systems (22). Spatial data can be easily incorporated into small-area level studies using a form of spatial data known as areal data (22, 27). With areal data, a chosen study region is divided into a set number of geographical units (areal units), where data points that fall within the boundaries of a unit are aggregated to that areal unit. Using confirmatory spatial analysis methods such as spatial regressions, the association between an outcome of interest and explanatory variables can be explored while accounting for the effects of geography across the study region (20). GIS and spatial analysis methods also provide an effective and simple way of visualizing data and possible spatial patterns by displaying data with maps (20, 22, 27). Furthermore, spatial cluster analysis methods can be applied to help determine the locations within the study region where high levels of alcohol expenditure exist (28), highlighting areas of interest that can serve as relevant targets for future public health interventions. Therefore, small-area level studies provide a useful way to incorporate spatial analysis methods into ecological studies and provide practical avenues for further action.

The City of Toronto serves as an ideal geographical area for an ecological study of alcohol expenditure using spatial analysis and GIS methods. The alcohol laws and policies that govern the use and distribution of alcohol are all uniform within the study region, as all alcohol retail policy is dictated by the Alcohol and Gaming Commission of Ontario (AGCO) (29). The large population of the City of Toronto provides an appropriate sample size to ensure the reliability and stability of the results from small-area analysis. In addition, the City of Toronto contains a diverse population in whom considerable variation exists for any socio-economic factor. As such, any differences in the effects of socio-economics on alcohol expenditure at the population level should be readily detected. Moreover, as the City of Toronto covers a large geographical area with many historical and distinct neighbourhoods, inherent spatial patterns are expected. The number of physical alcohol outlets in the study region is also expected to be high, in order to service a large metropolitan population, and therefore raises questions concerning the effects of alcohol availability on alcohol expenditure in the City of Toronto. Using GIS methods and spatial analysis allows public health officials to examine the associations between alcohol outlet density and alcohol expenditure

A more complete understanding of alcohol expenditure can help inform ways to reduce harmful alcohol use. In order to better understand the effects of geography, the built-environment and area level socioeconomic factors on alcohol expenditures, this study examines the associations between socio-economic and built-environment factors with alcohol expenditure at a small-area level in the City of Toronto. Furthermore, this study provides further insight by utilizing GIS methods to quantify and visualize the geographic patterning of alcohol expenditure in the City of Toronto.

## 1.1 Alcohol Expenditure Research Questions and Hypotheses

The purpose of this study was to explore the geographic distribution of three different types of household alcohol expenditures in the City of Toronto at a small-area area level in 2010. The study also examined the association between alcohol expenditure and socio-economic and built-environmental factors (alcohol outlets, commercial and residence use of land). My research questions were created to guide the study towards achieving these objectives. These research questions were applied to two types of average total household alcohol expenditures: **1) Average Total Household Alcohol Expenditure from Purchases in Stores (i.e. LCBO, Beer Store): 2) Average Total Household Alcohol Expenditure from Purchases in Licensed Premises (i.e. restaurants, bars, lounges). As a result, there are six research questions.**

For Average Total Household Alcohol Expenditure from *Purchases in Licensed Premises* the three research questions are:

**Question 1:** Is there geographic variation of the Average Total Household Alcohol Expenditure from *Purchases in Licensed Premises* in the City of Toronto? Furthermore, if geographic variation exists, where are the clusters of hot and cold spots?

**Question 2:** What socio-economic and built-environment factors (alcohol outlets, licensed premises, commercial and residential use of land, subways and highways) are associated with average total alcohol household expenditure by small-areas in the city of Toronto?

**Question 3:** Is there any remaining geographical variation in the average total household alcohol expenditure by small-areas in the City of Toronto after accounting for socio-economic and built-environment determinants?

For Average Total Household Alcohol Expenditure from *Purchases in Stores* the three research questions are:

**Question 4:** Is there geographic variation of the Average Total Household Alcohol Expenditure from *Purchases in Stores* in the City of Toronto? Furthermore, if geographic variation exists, where are the clusters of hot and cold spots?

**Question 5:** What socio-economic and built-environment factors (alcohol outlets, licensed premises, commercial and residential use of land, subways and highways) are associated with average total alcohol household expenditure by small-areas in the city of Toronto?

**Question 6:** Is there any remaining geographical variation in the average total household alcohol expenditure by small-areas in the City of Toronto after accounting for socio-economic and built-environment determinants?

# Chapter 2: Literature Review

Before performing a small-area level analysis of alcohol expenditure, it was important to understand the previous work done to study alcohol expenditure in order to guide the current study. The following chapter provides a review of alcohol expenditure data, the use of ecological models and applying spatial analysis in ecological models to study alcohol expenditure. The chapter will cover uses of alcohol expenditure data, previous studies of alcohol expenditure, individual versus ecological models, the social ecological model, and socio-economic and built-environment determinants of alcohol expenditure at the small-area level.

## 2.1 Alcohol Expenditure Data in Alcohol Use Analysis

Alcohol use can be measured either by the level of alcohol consumption or the level of alcohol expenditure. With alcohol consumption, alcohol use is measured by volume (i.e. ounces), primarily in the number of drinks consumed by the individual and is therefore usually measured as a categorical variable (i.e. one drink per day, two drinks per day, and more than two drinks per day). With alcohol expenditure, the total monetary amount spent on alcoholic beverages is recorded. Alcohol consumption has been the primary measure of study in examining alcohol use, as a large number of individual-level studies have examined the associations between differing levels of consumption and various alcohol-related harms. The results of these studies have consistently shown that high levels of alcohol use lead to the greatest risk for experiencing both acute and chronic alcohol-related illness (9, 30). As a result, the large majority of studies have sought to understand the determinants of alcohol consumption.

Due to the focus on alcohol consumption, studies on alcohol expenditure and in particular ecological studies of alcohol expenditure have remained largely unexplored. While alcohol expenditure is expected to be correlated to alcohol consumption, the use of expenditure data presents a different perspective; the use of alcohol expenditure also carries certain advantages over the use of alcohol consumption data. The first advantage is that alcohol expenditure data from sales data represents verified spending habits on alcohol products by individuals (16). This is particularly important as alcohol consumption survey data is known to suffer inaccuracies from issues of underreporting. This issue is further complicated by different levels of underreporting between various sub-groups such as in the case of heavy drinkers and moderate drinkers (31).

A second advantage is that alcohol expenditure data that comes from household surveys are also reflective of choices made by the survey respondent; therefore, the expenditure data incorporates influences from lifestyles, personal preferences and local market characteristics (16). In addition, alcohol expenditure data can be examined in relation with household income. A small-area level study by Grubesic, Pridemore, Williams et al. (16) controlled for alcohol expenditure by calculating alcohol expenditure as a percentage of median household income at the block group level. Other economic studies looked at determining demand equations for alcohol expenditure and have also indicated that alcohol expenditures as a proportion of household income holds significance (15, 32).

A third advantage is that alcohol expenditure can also be used to calculate the overall economic impact of alcohol use on a region. In Ontario, the sale of alcoholic beverages can only be done through authorized retail outlets such as restaurants, bars and liquor stores and represents a significant source of income for the province. Alcohol expenditures can therefore be used to examine the possible taxation and economic consequences for the region.

One further advantage when using alcohol expenditure data for ecological studies is the high level of availability of aggregated expenditure data (16). Alcohol expenditure data is widely available at very fine geographical scales, such as census Dissemination Area (DA) level making the data amenable to small-area level studies.

A possible concern with alcohol expenditure, however, is the accuracy with which the data coincides with alcohol consumption. Comparison studies, however, have shown that alcohol expenditure decisions are expected to coincide with consumption decisions (19, 33). Furthermore, at fine geographical scales the alcohol expenditure data is still reflective of the variations in alcohol-related commerce in the region, irrespective of how accurate the expenditure data reflects consumption or sales data (16).

## 2.2 Previous Studies of Alcohol Expenditure

Three previous studies have investigated the link between alcohol expenditure and socio-economic determinants (15, 32, 33). In all three quantitative studies, the expenditure data was retrieved from national surveys of household expenditure (15, 33), two of which were performed in North America while the other was performed in Asia. The primary focus of these three studies was to understand the causes of household alcohol consumption and levels of consumption at the individual-level using household expenditure. Therefore, a variety of economic statistical models were utilized, including the Double-Hurdle Model (15), Heckman's Sample Selection (32) and Gamma-Tobit model (34). While all three models take different approaches, they all reach the same objective in their results by producing two different equations to model the probability of alcohol consumption and the amount of consumption.

The U.S. study by Yen and Jensen (33) used individual-level household expenditure data and detailed socio-economic data from 1989 and 1990 retrieved from the Bureau of Labour Statistics (BLS) Consumer Expenditure Diary Survey. Yen and Jensen then used econometric statistical models to determine the maximum likelihood of ten different explanatory variables (income, regions of residence, education, marital status, gender and home ownership) affecting the probability of acquiring alcoholic beverages and the level of consumption if an individual decides to consume (i.e. Tobit statistical model and Double-Hurdle Model) (33). Results from the final analysis showed that income, region and household demographics were significant factors influencing the level of household alcohol expenditure (33). This study by Yen and Jensen, however, did not account for the possibility of broader contextual factors such as the surrounding environment or the built-environment (33). An example of this, would be alcohol retailers which operate at a small-area level and could potentially affect alcohol expenditure. Additionally, while the study had a variable controlling for region of residence, the variable used broad categories with unspecified geographic boundaries (33).

Tan, Yen and Nayga (32) used a Heckman Sample Selection to examine socio-economic factors that influence the purchase and amount spent on alcohol using data retrieved from the Malaysian Household Expenditure Survey from the 2004/2005 year. Similar to the study conducted by Yen and Jensen, individual-level factors were considered while geography and environmental factors were unaccounted for (32). Tan, Yen and Nayga reported that household expenditure on alcohol was significantly affected by a variety of socio-economic variables, including education, income, age, white-collar occupation and urban environments (32). The authors also considered the effect of socio-economics on household alcohol expenditure by different ethnicities (32). Results revealed that the association between certain socio-economic factors and household expenditure on alcohol varied by ethnicity (32). For example, a higher level of education decreased alcohol expenditures for Chinese households; however, education was insignificant for the Indian and Others ethnic group (32).

The other North American study by Abdel-Ghany and Silver (15) was the only one that examined the association between socio-economic determinants and household expenditure on alcohol in a Canadian context. The researchers looked at various socio-economic determinants for Canadian household expenditures on alcoholic beverages using data from the 1990 Survey of Family Expenditures published by Statistics Canada in 1992 (15). Significant findings included household tobacco expenditure, higher income, higher education levels, being male and age, all of which had positive effects on alcohol expenditures (15). Similar to the study by Yen and Jensen, while their work highlighted important household influences leading to positive alcohol-expenditure, broader contexts of environment, geography and locational factors were not taken into consideration (15). The authors did give some considerations to location within Canada by examining the province of origin; however, due to the large geographical boundaries of the provinces, smaller, local patterns of alcohol expenditure could not be examined (15).

## 2.3 Individual versus Ecological Approaches

Although all three studies provide an in-depth examination of the possible household factors related to alcohol expenditures, ecological factors were not considered in the analysis. Individual-level models, which are focused on proximal causes of health such as lifestyle and health behaviours, do not consider influences from the population level and therefore ignore potentially significant environmental factors. Modern epidemiology has often treated populations as aggregates of individuals; however, this may be an invalid assumption evidenced by the fact that certain population level health effects do not translate to the individual-level (25, 26). For example, at the population level, income inequality is negatively correlated with life expectancy (25, 26); however, there is no equivalent at the individual-level. Conversely, at the individual-level, life expectancy is positively correlated to individual income (26). If personal incomes are just aggregated into the population variable of average income, studies have found that there is little to no correlation between average income and life expectancy at the population level (26). In order to elucidate population level associations, ecological studies cane be used. Ecological studies are studies which use groups or populations of individuals as the unit of analysis as opposed to individuals (21). By focusing on populations, ecological studies bring about a different perspective as well as certain advantages over individual-level studies (21, 35). These advantages include a diversity of methods, compatibility with GIS and spatial methods, the ability to incorporate environmental factors and methodological advantages.

The use of ecological studies offers a great deal of flexibility as they are a part of a larger group of geographical variation studies which can be used in an exploratory fashion as well as in hypothesis testing (20, 35). The first type of ecological study is one that looks to describe the distribution of an outcome of interest (i.e. a disease) with respect to the place of occurrence (clusters of the disease in the study region) (21, 35). This first type of study is usually employed in an exploratory fashion in disease monitoring. The second type of study is a geographical correlation study, where the aim is to determine the association between geographical variations in an outcome measure to varying degrees of exposure to a certain factor (21, 35). Although correlations are most often carried out to quantify the associations, other confirmatory spatial analysis methods such as spatial regression can be employed for a better measure of the association (21, 35). Migrant studies, which examine if risk of disease for migrants change when they move from a high risk region to a low risk region (or vice versa) are also a type of ecological study (21, 35). Furthermore, as ecological studies can also include a time component, time-series and space-time studies can also be carried out for longitudinal studies (21, 35).

Geographical information systems and spatial analysis methods are a natural fit with ecological studies, as having a defined geographical region is important for ecological studies. Geographical information systems can help define geographical boundaries, map environmental variables and visualize data on maps to provide a succinct and impactful way of displaying the data (28, 36). Spatial analysis methods can also be used to help account for the geographical variation that is present in the data (20, 23, 37).

Ecological studies are also able to examine both socio-economic and environmental factors by taking into consideration the features of the location for the study. Environmental variables can be used to measure physical characteristics of an area (i.e. number of alcohol outlets). Socio-economic variables can be examined with aggregate measures that are made from individual measures grouped together to the area level (i.e. percentage of heavy drinkers). For this study, the environmental variables are alcohol outlets, commercial and residential use of land, subways and highways.

Often times, data limitations also make ecological studies the only viable route. Individual-level data for a certain exposure of interest is frequently unavailable as it is may not be practical to collect or doing so would be cost-prohibitive (21). Using ecological studies, exposure can be examined from an area level, avoiding uncertainties about individual exposure. Socio-economic and other common statistics are also collected by area level on a routine basis making ecological analysis possible (21). Due to this availability in data, ecological studies can often be accomplished quickly and inexpensively.

Ecological studies can also take advantage of large differences in exposures between different areas to help find associations that may otherwise be hidden at the individual-level (21, 35). Finally, the random errors for exposure are often smaller for populations than for individuals, making ecological studies more attractive, as random errors are known to cause regression-dilution bias in the analysis (21, 35).

## 2.4 Applying Spatial Methods to Ecological Studies

Spatial methods or spatial analysis is the ability to make use of spatially referenced data by processing it into different forms to acquire additional information and meaning from the results. To do this, a specialized set of tools known as geographical information systems (GIS) is used (22). GIS are automated systems that are designed to specifically capture, store, retrieve, analyze and display spatially referenced data. The uniqueness of a GIS is that not only is it a database but it also contains a map component or spatial information which is associated to each piece of data. Therefore, within a GIS database there are two different types of data stored: attribute data and map data (22, 38). For this study, area level data was the type of spatial data that was used.

Using area level analysis, the area of analysis can be broken down into geographical units or small-area units that have comparable population counts between each area. This type of small-area analysis is able take into account how alcohol expenditure data is influenced by the chosen explanatory variables in each of the areas in the study region. Small-area level data is sensitive enough to reveal meaningful local patterns (22) that would have been otherwise masked at much broader scales (i.e. using provinces). Spatial analysis methods such as clustering methods and spatial regressions can then be performed to look for geographical patterns and to confirm associations between alcohol expenditure data and various ecological factors.

Finally, using spatial methods provides a powerful way to display, visualize and present the results of the analysis through maps. To date, there have been no small-area level studies utilizing alcohol expenditure as the main outcome variable to examine the influences of socio-economic factors on alcohol expenditures. This study used geographic coordinates to define specific regions and boundaries in order to make use of GIS and spatial analysis methods to explore geographical patterns.

## 2.5 Determinants of Alcohol Expenditures

### 2.5.1 The Social-Ecological Model

In this study, the social-ecological model was used as a guiding framework in examining small-area level determinants of alcohol expenditure. The general social-ecological model was first conceived and proposed by Urie Bronfenbrenner (39). Bronfenbrenner (39) proposed that human behaviour and development is affected by the entire ecological system around the individual. The influences of the environment can be separated into multiple levels going from proximal to distal environments (39). For this study, the version of the social-ecological model that will be used as a framework is the social-ecological health model outlined by Coreil (1) in Figure 2.5.1.1. This model arranges determinants of health in five different hierarchical categories, beginning with individual characteristics at the intrapersonal level, followed by the interpersonal, organizational, community and society levels respectively. Health consequences are understood to be influenced by all five levels, with each level intertwining with all of the others. As a result, good health is a not only a matter of personal behaviours, but is also heavily influenced by the social and physical environment (40). For the current study, the social-ecological health model was used as a guiding framework to examine the area level characteristics associated with alcohol expenditure. In the case of this project, the socio-economic and environmental factors population at the small-area level represent influences that lie within the community layer.

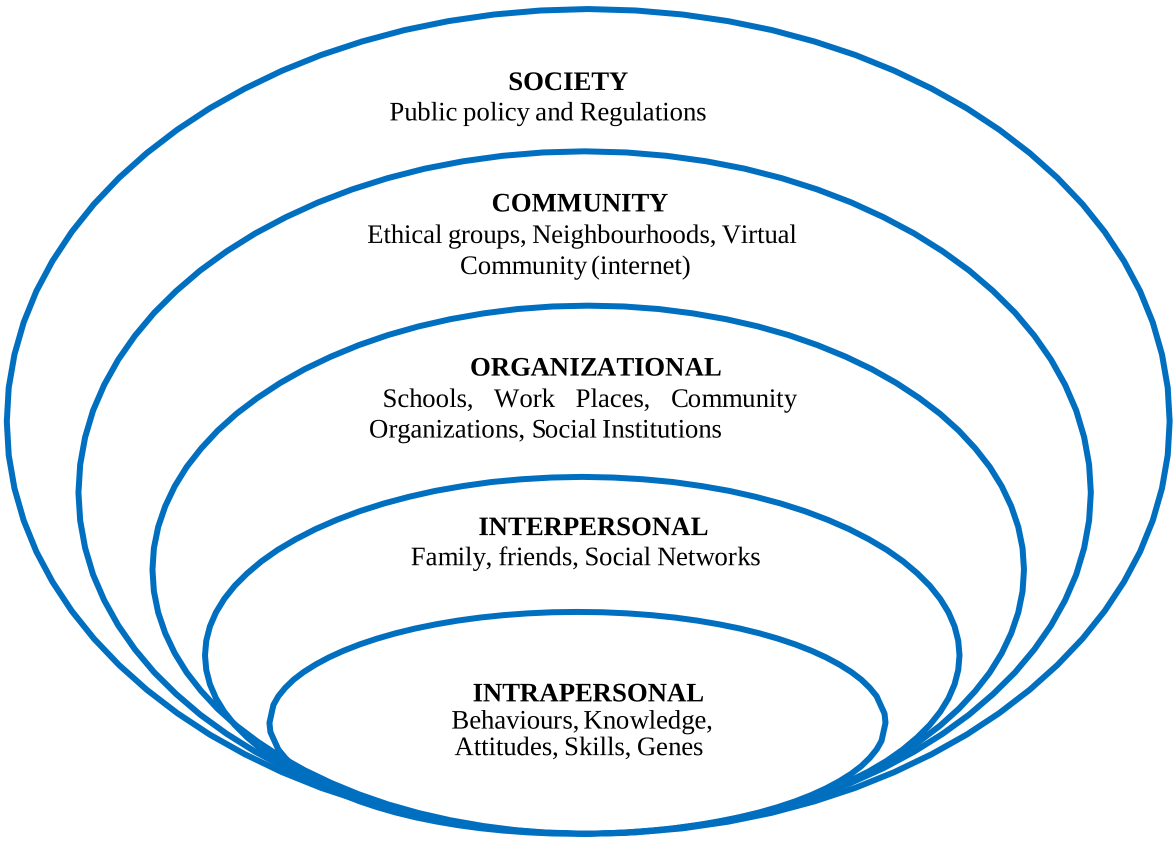



Figure 2.5.1.1: The social-ecological health model (1)

### 2.5.2 Ecological Factors for Alcohol Expenditure

The two set of ecological factors that examined in the study are socio-economic variables at the small-area level and built-environmental factors. The small-area level determinants of alcohol expenditure that were included in the analysis of this study were determined through a literature review of both the individual-level expenditure studies mentioned before and studies of alcohol consumption. Studies of consumption globally have shown that the level of drinking varies greatly by geography because of social factors, economic development, cultural influences, and policies regarding alcohol outlets. Even within a given country, strong regional differences exist with consumption as these various socio-economic factors can vary greatly on a regional basis.

The built-environment is defined as the human-made environment that is possibly subject to planning and refers to all physical environments and elements including buildings, infrastructure and green space that is constructed for and by human activity (24, 41). Aside from the physical elements, the built-environment also consists of processes such as urban design, land development, transportation development and patterns of human activity (24, 42-44). These processes are also influenced by economics, politics and environment of the region (45). There is a growing body of evidence that now supports the importance of the built-environment in either promoting or hindering the adoption of health behaviours, especially that of physical activity (24, 42, 45-48). Furthermore, the built-environment has been recognized for its significant association with chronic diseases (41, 42, 45, 48-50). The built-environment can also play significant role in influencing alcohol expenditures, as the number of alcohol outlets and the land set aside for other alcohol retailers can have a significant effect on the levels of alcohol consumption in the region.

For this study, eleven categories of socio-economic factors at the small-area level and six built-environment factors were examined. The following section elaborates on the socio-economic and built-environment determinants of alcohol expenditure that will be included into the analysis to help answer research question three.

### 2.5.3 Socio-Economic Factors

*Age*

For high income countries, the highest levels of consumption generally occur among young people in early adulthood whereas in middle to lower income countries, the largest amounts of alcohol are consumed during midlife (ages 35 – 60) (3). Canada fits the profile for high income countries as heavy drinking is most prevalent between the ages of 18 to 34 (51, 52). When considering all forms of drinking that exceed the low-risk drinking guidelines, a larger proportion of those aged 15 to 24 exceed the Low-Risk Drinking Guidelines (LRDG) for both acute and chronic consumption, as compared to the proportion of those who do so in the 25 and above age category (53).

*Gender*

There are significant gender differences in alcohol consumption that have persisted both historically and cross-culturally (3). Primarily, men have always consumed alcohol in greater quantities than women. Women also have higher rates of abstention from alcohol as compared to men, while men are much more likely to consume alcohol at problematic levels (3, 54). In Canada, heavy drinking rates reflect similar gender differences, with considerably more men participating in heavy drinking as compared to women across all age groups (51, 52).

*Ethnicity*

Individual-level studies in Malaysia, US and Korea have all highlighted the significant effects of ethnicity on alcohol consumption. In Malaysia, Tan, Yen & Nagaya (32) showed that Chinese consumers were more likely to spend less on alcohol as result of higher education whereas the effect of higher education on consumption for Indian and other ethnicities[2] was not significant. In a study of Korean alcohol consumption, Sharpe et al. (19) found that gender and education had the same influence on consumption in Korea as well as in North America, but effects of other factors such as household economics and family structure differed between the two locations. Studies in the US and Canada (15, 17, 18, 55, 56), also indicate that Caucasian people and Caucasian-headed households are more likely to consume alcohol in comparison with minorities and minority-headed households respectively.

*Occupation*

A number of individual-level studies have also shown that occupation has a significant impact on alcohol consumption. Sharpe et al. (19) found that housewives had the lowest likelihood of consuming alcohol, while those who were self-employed and worked in the farm and fisheries industry consumed the most amount of alcohol. Berry et al. (57) found that in Australia, alcohol consumption was greater among blue-collar workers than those in white-collar jobs. Pomerleau, Pederson and Østbye et al. also found that in Ontario, higher occupational prestige was associated with the lowest odds for alcohol consumption. A similar study in Malaysia showed a different association, as Tan et al. (32) found that households with white-collar occupations were more likely to purchase alcohol in comparison to blue-collar households. In Canada, Abdel-Ghany and Silver (15) found that teaching was the only occupation associated with higher levels of alcohol consumption. Further studies are needed to clarify the association between alcohol expenditure and socio-economic factors.

[2] Ethnicities included Malay, Chinese, Indian and Other

*Income*

While there have been studies that have suggested neutral association with alcohol consumption (19), the majority of studies have consistently pointed to a positive association between income and alcohol consumption (15, 17, 18, 32, 56, 58). In Ontario, Pomerleau, Pederson and Østbye et al. also found that higher income was positively associated with alcohol consumption. Possible reasons for this association include alcohol being relatively less costly in high income countries. In addition, households with higher incomes are assumed to have higher levels of disposable income and therefore, can afford to consume greater amounts of alcohol.

*Housing*

A high proportion of rented dwellings in an area is commonly an indicator of income and housing instability (59, 60) and also serves as a measure of material deprivation[3]. Previous studies examining the association between material deprivation and alcohol consumption have produced mixed results (63-67). Housing measures were included in the analysis to explore possible associations with expenditure levels.

*Marital Status*

Studies in the US, Australia and Korea have found that marital status is significantly related to alcohol consumption, with married couples being less likely to consume alcohol (19, 33, 57). Single persons were more like to drink alcohol at every type to consumption level, including risky drinking, infrequent and occasional drinking (57).

[3] Material Deprivation is the inability for individuals or households to afford basic goods, needs (i.e. food with meat, safe housing, etc.), and is used to provide an additional perspective on poverty. (61, 62)

*Lone-Parent Households*[4]

Examining lone-parent headed households provides a measure of family disruption which also provides another measure of poverty and material deprivation. Generally, lone-parent households are more prone to lower socioeconomic status, which is negatively associated with alcohol expenditure (61, 69-71). Furthermore, having dependents makes it increasingly unlikely that the household would spend money on alcohol, leading to a negative association between number of children and alcohol expenditure (32, 33, 72).

*Education*

In Western nations such as Canada, higher levels of education have been found to positively associate with both frequent and heavy drinking and the likelihood to consume alcohol (33, 73). Pan et al. proposed that higher educational attainment provided for a greater opportunity for social interactions which lead to more opportunities to consume alcohol (58). However, there are some conflicting results as higher education has also been demonstrated to decrease the risk of heavy and risky drinking in workers such as in Australia, Korea and Malaysia (19, 32, 57).

*Average Number of Persons in a Private Household*

Previous studies (17, 32, 33, 72) have also demonstrated that as the number of individuals in the family increases, the alcohol expenditures by the family tends to decrease. Tan's (32) explanation for this association was that alcoholic beverages are not a basic necessity and expenditure on alcohol would decrease in proportion to the increase in average persons in a household.

[4] Lone-parent households are a type of family structure with a lone male or female parent with dependents (68)

*Aboriginal Identity*

Aboriginal peoples represent an important and distinct minority group that also need to be considered separately (74). Compared with the rest of the Canadian population, the aboriginal population have suffered from a greater level of social inequality and poorer social determinants of health (75). As a result, they have faced increased mortality and morbidity from chronic diseases, mental issues and substance abuse, including alcoholism. Alcoholism remains a prevalent issue within the Aboriginal community. Studies of aboriginal youth have also showed a high prevalence of binge drinking. In Ontario, the off-reserve aboriginal peoples have higher rates of exceeding the LRDG than the general public (76). For off-reserve First Nation's peoples, 12% of adult males exceed the LRDG compared with only 8.4% in non-aboriginal adults. By comparison, rates of aboriginal female adults and non-aboriginal female adults that exceed the LRDG was similar at around 7% and 7.7% respectively (77). For the Métis population, both the male and female populations exceeded the LRDG over their non-aboriginal counterparts at 11% and 10% respectively (77).

### 2.5.4 Built-Environment Factors

*Land Use*

Land Use refers to the classification of an area of land by the human activities that occur within that area (78). Common categories of land use include residential, industrial and commercial areas. Land use is a function of government zoning regulation and can be subject to change over time (16, 79). Even between adjoining regions, the types and proportions of land use in each region are unlikely to be homogenous, as they will probably contain their own unique blend of land uses. As land use dictates the types of activities that occur in an area, it can affect the volume of human traffic, the business and industries that reside there as well as the recreational activities that take place (16). For this study, residential use of land and commercial use of land were examined.

### *Residential Use of Land*

In studies examining the association between violence and alcohol outlets, certain types of land use were found by Snowden and Pridemore (80) to directly affect violence. A study by Pridemore and Grubesic (81) found that in urban centres, block groups with higher proportions of single-family residential use areas or commercial use areas resulted in a weakened association between outlet density and assaultive violence. This indicates that residential and commercial use of land may mitigate the consumption of alcohol in some way, thereby weakening the connection between increased alcohol outlet density and assaultive violence. It is possible that expenditures are negatively impacted by residential use of land. As such, residential and commercial use of land is a potential factor in determining alcohol expenditures at the small-area level. The urban environment has also been associated with higher levels of consumption (32) and researchers have found that urban environments have been targeted by alcohol and tobacco marketers (49, 82). As a result, areas with denser population may be associated with greater alcohol expenditures.

### *Commercial Use of Land*

A strong body of evidence now exists that links even small increases in the number of alcohol outlets to increased consumption (83). This in turn leads to a positive association between alcohol outlets and alcohol-related harms. Studies have shown that these increases in the availability of alcohol leads to increased sales of alcohol independent of price increases (83).

More recent studies using geographical units have also shown that alcohol-related harms are positively associated to alcohol outlets at the census tract level and also at even smaller scales (83). The findings suggest that small-area level analysis using the smallest geographical units can help reveal important associations regarding alcohol outlets and alcohol expenditure that are otherwise hidden at larger geographical units. Furthermore, these findings also point to the importance of how land is zoned locally by communities. By controlling land use, local governments have successfully been able to regulate the number of alcohol outlets to minimize alcohol-related harms (83). As a result, the amount of land in an area that is designated for commercial use will affect the number of alcohol outlets that are allowed in the region. Therefore, the proportion of commercial land in an area was expected to have an effect on the alcohol expenditure levels.

*Alcohol Outlet Densities*

To date, alcohol outlets has been the most widely studied variable using spatial analysis methods (84). One of the main reasons for this is that outlets are all associated with a geographical location making them immediately compatible with GIS and spatial analyses. Much of the research on availability has involved helping elucidate the association between outlet density and alcohol-related harms. In particular, some of the most widely studied and established associations are between outlet density and motor vehicle crashes (85). Outlet density has also been examined in the context of consumption and was found to be a significant factor (84).

Only a few ecological studies have examined alcohol outlets and expenditure, in large part due to a lack of alcohol outlet and expenditure data. A study in Norway found a positive association between alcohol outlet density and alcohol sales[5] (91, 92). Other international studies of alcohol availability also mirror these findings. Studies in the United States have provided mixed conclusions: only one state-level study provided evidence of a positive association between alcohol outlet density and alcohol sales (93). However, data limitations prevented replication of the results. Survey data used to examine the same association at the individual-level has also provided mixed results (60).

At the individual-level, alcohol outlets may also moderate the association between socio-economic variables and expenditure. Certain studies have shown that greater availability is associated with low income areas, high unemployment, crowded housing and areas with low rent pricing (60, 94). This association has also been demonstrated at the small-area level, and as a result, alcohol outlets should be controlled for in the analysis (60).

Additionally, associations have been established between alcohol outlets and alcohol-related harms (95, 96) and in particular with violence (86-90). A large number of population level studies have established a positive association between outlet density and a variety of alcohol-related harms. These include traffic accidents, various crimes and violence (91). If a significant association is found between expenditure and alcohol outlets, it may also mean that alcohol-related crimes may be similarly related to expenditure. Future analysis could also include an examination of expenditure and crime; for these reasons, alcohol outlets figured to have a significant effect on expenditures and were included in the analysis.

[5] Alcohol Outlets also found to be positively correlated to violence (86-90)

*Restaurant Density*

Alcohol outlets can be separated into two categories: off-premise and on-premises outlets. Off-premises outlets are stores where purchased alcohol is not consumed at the store. Examples of off-premises outlets include liquor stores, grocery stores and convenience stores. On-premises outlets consist of bars and restaurants. For this study, restaurant density and primary drinking restaurant density referred to on-premise facilities and outlet density referred to off-premise facilities

*Subways and Highways*

Subways and highways are expected to increase alcohol expenditures by increasing access to alcohol outlets. Pollack, Cubbin, Ahn and Winkleby(60) suggested that public transit and the use of cars may extend the distances people are willing to travel in order to purchase alcohol. For lower density areas such as suburbs, where people are more likely to have automobiles, highways may enable greater access to alcohol outlets in neighbouring areas.

### 2.5.5 Summary

This literature review demonstrates the need to study expenditure as well as highlights the absence of small-area level studies. Furthermore, population level studies provide a broad scope by enabling the examination of environmental factors. Spatial analysis methods also provide the ability to quantitatively incorporate geography into the study and to help determine the geographical variation of expenditure at the small-area level. The potential findings between expenditure and population level factors in this study can be combined with individual-level determinants of expenditure, resulting in a more complete understanding of what drives spending on alcohol. These findings can then be used to guide future policies looking to curb harmful alcohol use.

## 2.6 Study Rationale

Alcohol consumption remains a significant contributor to morbidity in Canada through both chronic disease and acute harms and injuries. As a result, aggregate harmful consumption leads to a heavy burden on the health care system. Despite these alcohol-related harms being well documented, a significant proportion of Canadians continue to report drinking above the LRDG levels. Among drinkers, 18.7% of Canadians reported levels of consumption exceeding the LRDG for chronic effects and 13.1% of Canadians reported consumption exceeding the LRDG for acute harms in 2011 (97). A significant proportion of Canadians have also reported engaging in heavy drinking which carries the greatest potential risk of alcohol harms. In Ontario, LRDG drinking patterns among drinkers have mirrored national trends, with 18.8% of drinking Ontarians reporting levels of consumption exceeding the LRDG for chronic effects and 12.8% reporting drinking above the LRDG for acute harms (98). To help elucidate the contributing factors to these high levels of use, the existing literature has focused on examining factors related to consumption at the individual-level. Consumption data, however, is not without issues as there are problems with underreporting. As a result, consumption data may not be as effective at determining the economic impact of use at the population level. Furthermore, individual-level models do not account for possible environmental influences. Examining alcohol expenditures at using an ecological study can help to remedy these problems. Ecological studies are able to incorporate environmental variables into the analysis as they use the area level as the unit of analysis. Using ecological studies also naturally allows the use of GIS and spatial methods to help explore and map the distribution of alcohol expenditure in the City of Toronto.

Additionally, confirmatory spatial analysis methods such as spatial regression can be used to help account for geographical variation while examining the association between alcohol expenditure and various explanatory variables. The resulting spatial regression model can also provide the ability to predict changes in alcohol expenditure values from changes to the values of the explanatory variables. This study presents an alternative perspective by exploring the ecological factors that affect expenditure levels. More specifically, the current study aimed to explore the geographical distribution and to see if geographic variation exists in two types of alcohol expenditure data in 2010 at the Dissemination Area level. Furthermore, this study looks to identify socio-economic and built-environmental determinants of expenditure at the small-area level by utilizing spatial regression models.

# Chapter 3: Methods

For this study, spatial analysis and spatial regression methods were applied to identify geographic variation for alcohol expenditure data as well as to identify small-area level socio-economic and built-environment determinants of alcohol expenditure. This chapter begins by defining the geographical unit of analysis as well as providing geographical context for the City of Toronto. This is followed by a discussion of the data sources used for the study, detailing which socio-economic and built-environment variables were included in the study and how they were incorporated into the analysis. The second section details how global and local cluster methods were applied to answer research question one and four. The chapter concludes with the final section describing how spatial regressions were used to answer research questions two and three as well as research questions five and six.

## 3.1 Unit of Analysis, Geographical Context, Data Sources and Measures

### 3.1.1 Unit of Analysis

The unit of analysis is the Dissemination Area (DA), which is the smallest census geographical unit that covers all of Canada, with each area containing 400 to 700 persons (99). Compared with larger census geographical units, socio-economic and built-environment variables are more homogenous within each DA. Therefore, using such a small spatial scale allows this study to detect patterns that might otherwise be masked using larger geographical units due to the aggregating of smaller areas. Figure 3.1.2.1 shows the 3685 Dissemination Areas that cover the City of Toronto.

### 3.1.2 Geographical Context – City of Toronto

The City of Toronto is the most populous city in Canada and is the core of the Greater Toronto Area (GTA), the most populous Canadian Metropolitan Area (CMA[6]). Aside from the City of Toronto, the GTA consists of four other regions: Halton region, Peel Region, York Region and Durham Region. The current City of Toronto was also previously separated into six former cities. As of the latest census, in 2011, there are 2.6 million people that live in the City of Toronto and a total of 6 million people live within the Greater Toronto Area.

The approximate boundaries of the city are the Ontario 427 highway to the West, Steeles Avenue to the North, the Scarborough-Pickering Townline to the East, and Lake Ontario to the south. There are also several major highways across the city, specifically: Highways 400, 401, 404, the Don Valley Express Way (DVP), 409, 427 and the Gardiner Expressway (GE) and the Allen Road expressway. Figure 3.1.2.1 displays a map with only the major expressways while Figure 3.1.2.2 illustrates all of the expressways and major roads within the City of Toronto. Both the major highways and roads served as important landmarks that were used to help identify the locations of hot-spots (areas of high alcohol expenditure values) and cold-spots (areas with low alcohol expenditures values) found in this study.

[6] A CMA is an area consisting of one more or more municipalities located around a major urban core and must have a total population of at least 100,000 with 50,000 or more living within the urban core (100).

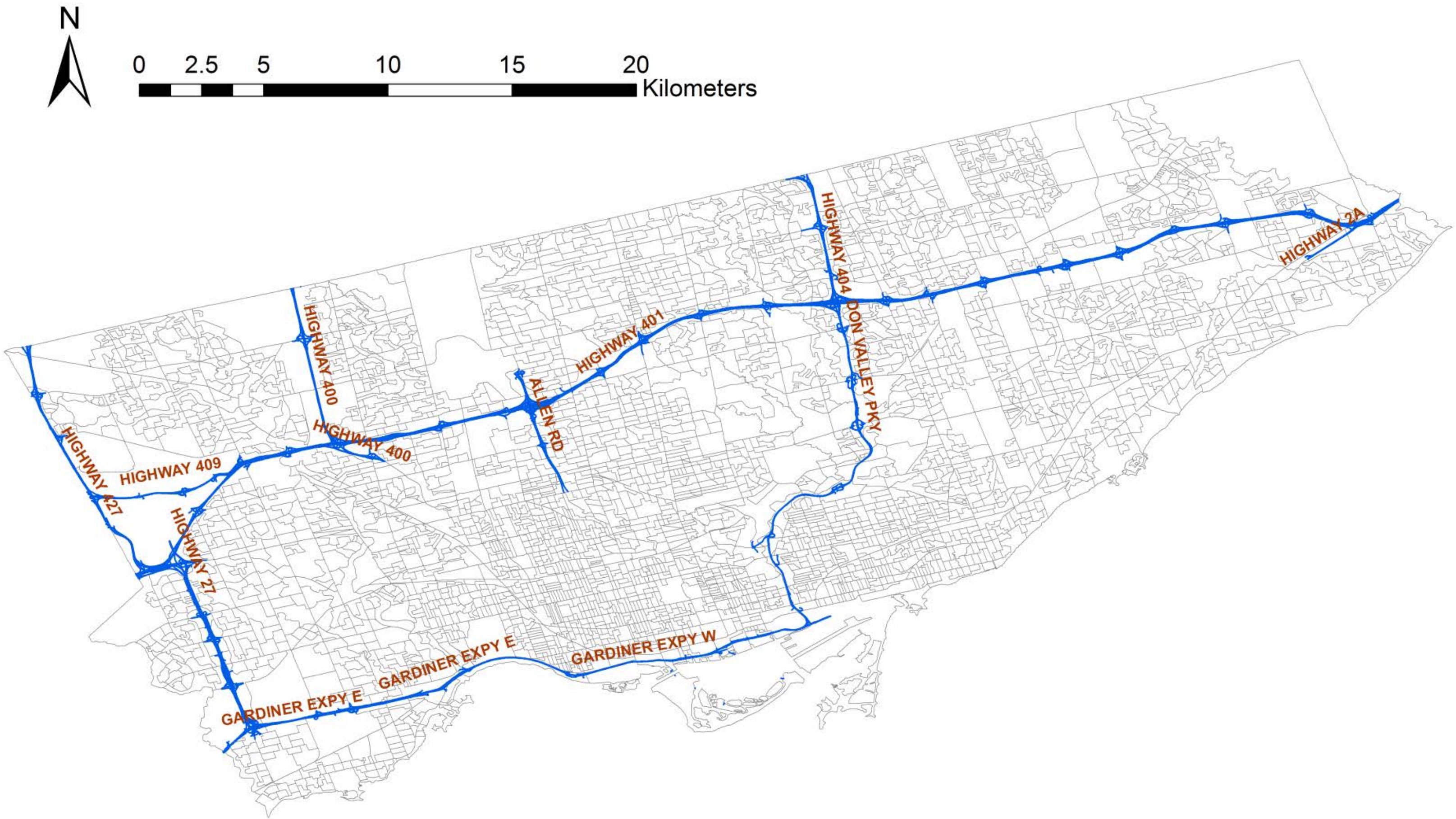



Figure 3.1.2.1 - Dissemination Area Map of Toronto with Major Highways

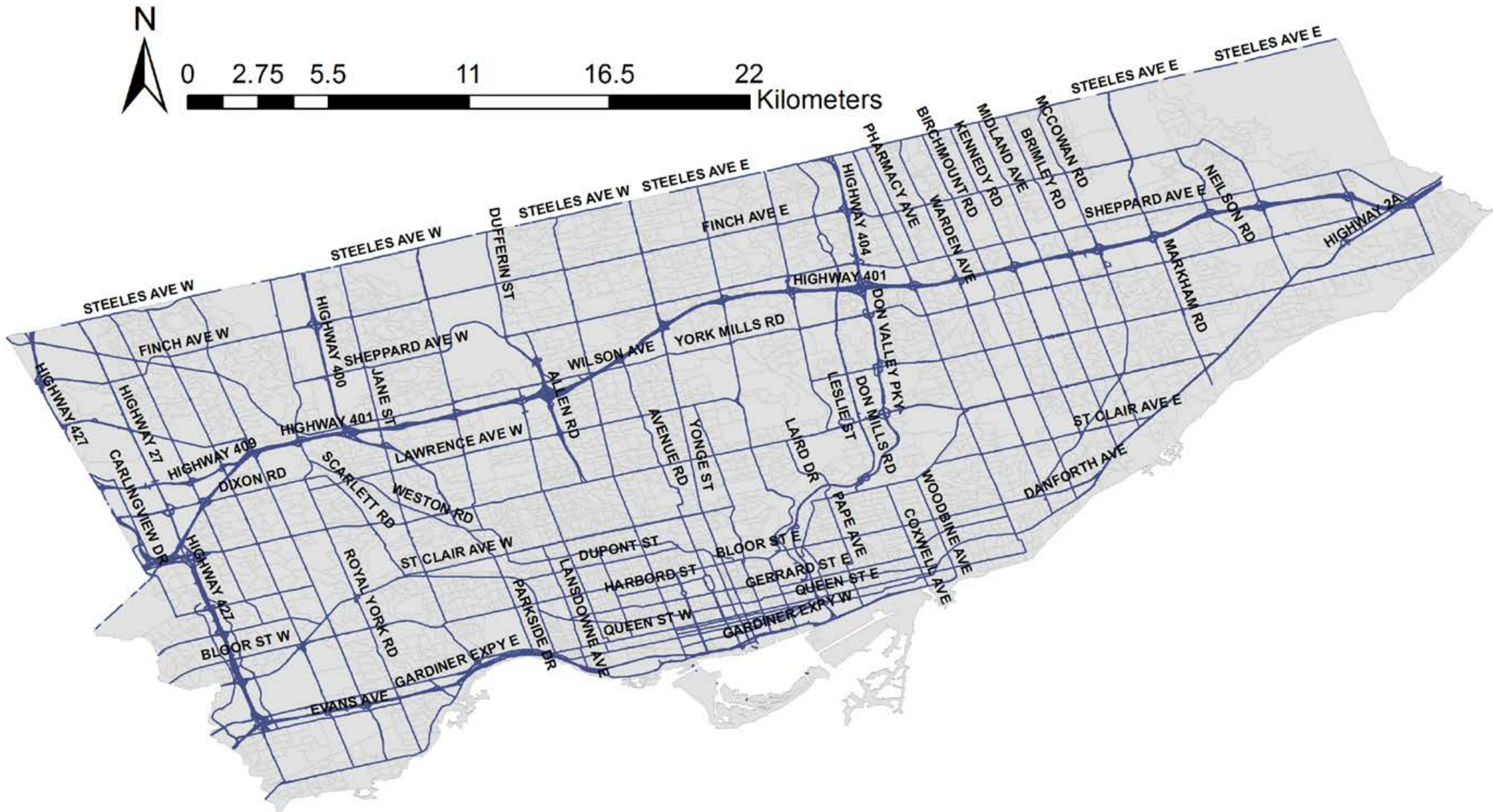



Figure 3.1.2.2 - Map of Toronto - Major Streets and Highways

Before the City of Toronto amalgamated in 1998, the current City of Toronto was divided into 6 different municipalities: Etobicoke, North York, Scarborough, York, East York and former Toronto (Figure 3.1.2.3). While these municipalities are now constituent parts of the City of Toronto, the names of these former municipalities are still used by residents who reside in those respective areas as they provide a quick geographic reference when discussing the locations within the City of Toronto. This study included these historical municipal names to describe the location of hot-spots and cold-spots, providing geographical context. York, East York and Toronto were referred to as downtown Toronto.

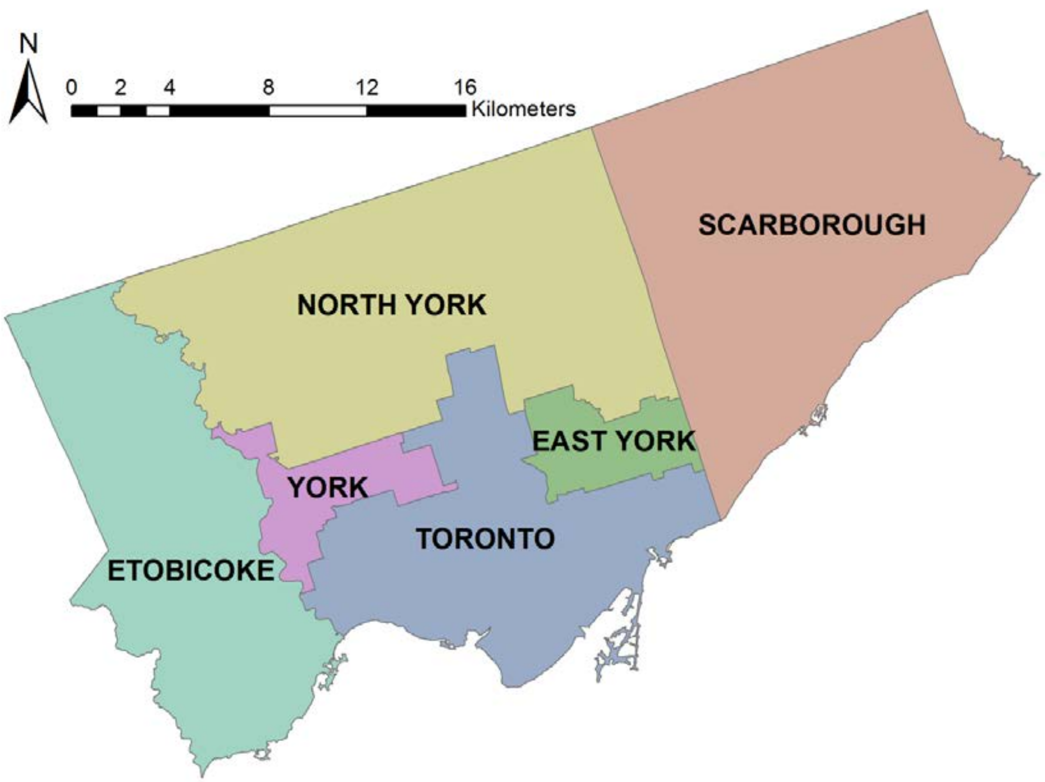



Figure 3.1.2.3 - Map of cities in Pre-amalgamated Toronto

Finally, there are a total of one hundred and forty recognized neighbourhoods within the City of Toronto (101). These neighbourhood boundaries were constructed by the City of Toronto from census tracts in order to create meaningful geographic areas (102). The boundaries of these neighbourhoods are also based on seven different criteria to help balance the population size of neighbourhoods, to respect existing boundaries for neighbourhoods, to make the neighbourhoods useful for service decisions and for data analysis purposes (101). Figure 3.1.2.4 displays a map of all one hundred and forty neighbourhoods in the City of Toronto. After the spatial distribution of alcohol expenditures was identified, GIS methods were used to correlate the Dissemination Areas to the neighbourhoods, helping to identify neighbourhoods with high levels of expenditure.

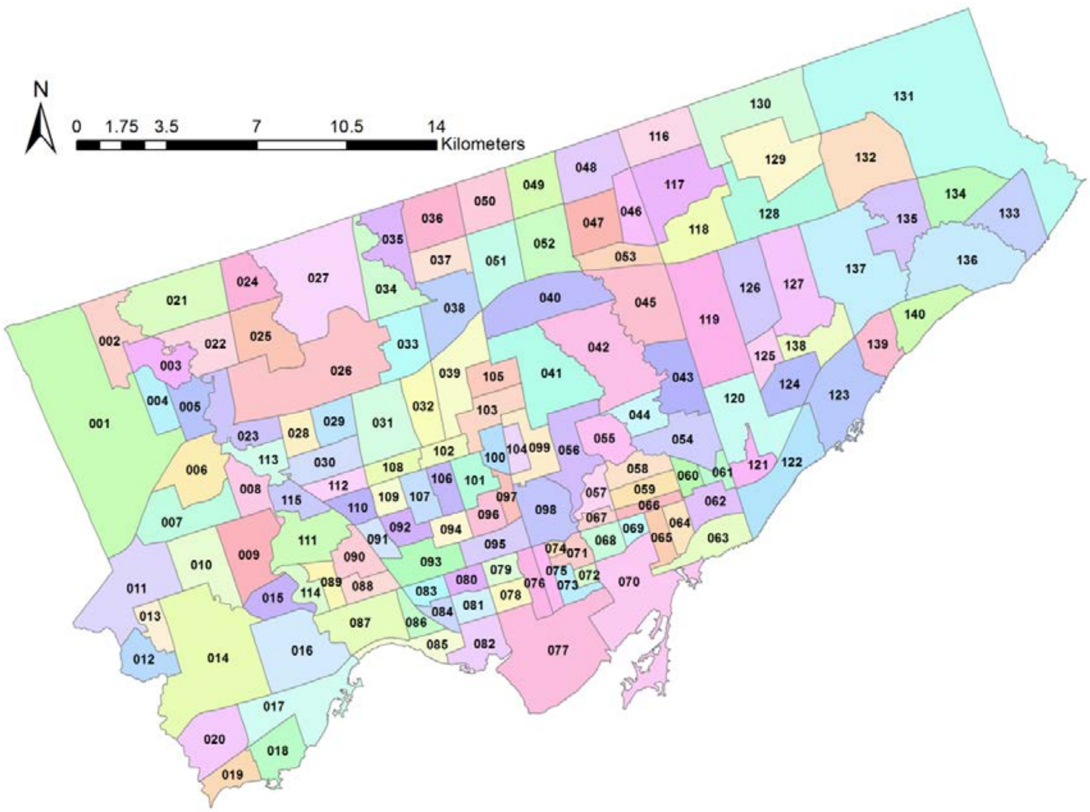



Figure 3.1.2.4 - Map of 140 Recognized Neighbourhoods in Toronto (**103**): The one hundred and forty recognized neighbourhoods are labeled with their neighbourhood number from 001 to 140.

### 3.1.3 Data Sources

Table 3.1.3-1 summarizes information on data sources that were used to gather the dependent and independent variables for this study. Information on the year of the dataset, description of the dataset, collection methods and variables the dataset provided are covered in the table.

Table 3.1.3-1: Summary of Data Sources

| Data Source | Year | Description |
|---|---|---|
| The Survey of Household Spending | 2010 | The Survey of Household Spending (104) is an annual cross-sectional survey that obtains detailed household expenditure profiles as well as information on dwelling characteristics and household equipment. The survey also collects data on annual incomes of household members as well as some demographic characteristics of the household (i.e. employment status). Dwelling characteristics such as the age, type and tenure of the household are also collected. Data collection is done through a Stratified Two Stage Plan. In stage one, geographical clusters are identified and in stage two, dwellings in those clusters are further identified to be candidates for the survey. The survey consists of a questionnaire that requests for expenditure information using reference periods of 1, 3 or 12 months depending on the expenditure type. The questionnaire is administered in person by an interviewer that has the survey on a laptop. A daily expenditure diary is also collected for a further sub-population of the sample.<br><br>Variables Provided:<br><br>1) Average total household alcohol expenditure from purchases in Stores;<br>2) Average total household alcohol expenditure from licensed premises; |
| 2006 Census of Canada | 2006 | The Canadian census (105) is a cross-sectional survey that is administered nationally to all households to collect vital data on the demographic, social and economic characteristics of all Canadians. For the 2006 Census of the Population, there were two versions: a short form and a long form census. The bulk of the demographic, social and economic data however, was collected through the Long Form of the Census that was sent to a representative sample that was consisted of 20% of the population. The 2006 Census also had a very high response rate of 96.5% due to the survey being mandatory.<br><br>Short Form of Census: Contained the basic data questions that asked age, sex, marital status, mother tongue and relationship to Person 1[7] (32.5 million persons or 12.7 million households) |

[7] Person 1 refers to the first person reported on the Census of Population Questionnaire (106)

| Data Source | Year | Description |
|---|---|---|
| | | Long Form of the Census (20% Sample): A regression estimation weighting procedure was employed to account for any under-represented groups. This weighting procedure known as calibration estimation was employed to extrapolate the data from the representative sample to 100% of the population.<br><br>Variables Provided:<br>1) Demographic, social and economic characteristics<br>2) Population counts<br>3) Map of Toronto using standardized small geographical units (Dissemination Areas) |
| Digital Mapping Technologies Inc. (DMTI) Spatial – Enhanced Points of Interest (EPOI) | 2008 | The 2008 DMTI – EPOI (107) dataset by DMTI Spatial is a database containing over 1 million business and recreational points of interest nationally across Canada. The database includes a large variety of built-environment structures, each individually geocoded to provide their location and building type. Examples include carpool parking lots, educational institutions and healthcare facilities. The following variables were retrieved from this dataset:<br><br>• Geo-coded locations of alcohol outlets, primary drinking restaurants and restaurants in Toronto |
| DMTI CanMap Route Logistics | 2010 | The DMTI CanMap Route Logistics (108) are a set of DMTI Spatial products that provide detailed road and highway maps of Canada as well as locations and detailed data for different types of infrastructure in each province. Specifically, the Land Use Layer Map for Ontario, a vector map which features seven land use categories, was used to provide the following land use data for the City of Toronto (78):<br><br>• Proportion of commercial land and residential land in the City of Toronto |
| City of Toronto Open Data – Data Catalogue: Neighbourhoods | 2014 | The shapefile with the digital boundaries of the City of Toronto neighbourhoods was retrieved from the City of Toronto Open Data – Data Catalogue under the Neighbourhoods section (109). |

### 3.1.4 Measures

Table 3.1.4-1 summarizes all of the expenditure outcome variables, socio-economic explanatory variables and built-environment explanatory variables that were used in the analysis.

Table 3.1.4-1: Description of All Measures

| Variable | Variable Type | Calculation of Variable |
|---|---|---|
| **Outcome Variables** | | |
| ***Average Total Household Alcohol Expenditure from purchases in Stores*** | C | The two variables were retrieved at the Dissemination Area level from the Survey of Household Spending through the SimplyMap database for the 2010 year. |
| ***Average Total Household Alcohol Expenditure from Licensed Premises*** | C | |
| **Ecological Determinants (Independent Variables)** | | |
| **Socio-Economic Variables** | | |
| ***Age*** | R | Percentage of 20 to 44 year olds age group. |
| ***Gender*** | R | Ratio of Males to Females |
| ***Ethnicity*** | R | Percentage White, Percentage Ethnicity South-Asian, Percentage Ethnicity Chinese, Percentage Ethnicity Black and Percentage Ethnicity Filipino. |
| ***Occupation*** | R | Percentage White-collar Occupations (management occupations; business, finance and administration (with secretaries/clerical worker category subtracted); natural and applied sciences; health and medicine; social sciences, education, government and religion)<br>Percentage Blue-collar Occupations (manufacturing, processing and utilities; construction; trades, transport and equipment operators) |
| ***Income*** | C | Median After-Tax Household Income |
| ***Social Assistance*** | R | Percentage of Income from Government Transfer Payments |
| ***Housing*** | R | Percentage of Apartments – 5 floors or more, Percentage of Apartments – less than 5 floors, percentage semi-detached, percentage fully-detached, Percentage of Row-Houses and Percentage of Rented Dwellings |
| ***Marital Status*** | R | Percentage Married |
| ***Lone Parent – Headed Households*** | R | Percentage of Lone Parents |
| ***Education*** | R | Percentage with Post-Secondary Education |

| Variable | Variable Type | Calculation of Variable |
|---|---|---|
| ***Average Number of Persons In a Private Household*** | C | Average number of persons in a private household |
| ***Aboriginal Identity*** | R | Percentage of aboriginal identity |
| **Built-Environment Variables** | | |
| ***Alcohol Outlet Density*** | R | By Area (# of Alcohol Outlets in DA/Area of DA), By Population (# of Alcohol Outlets in DA/Population in DA) |
| ***Residential Use of Land*** | R | The proportion of residential land to all land uses within a DA |
| ***Commercial Use of Land*** | R | The proportion of commercial land to all land uses within a DA |
| ***Subway Intercepts*** | B | Presence of Subway Line in DA = 1, No Presence of Subway Line in DA = 0 |
| ***Highway Intercepts*** | B | Presence of Highway Line in DA = 1, No Presence of Highway Line in DA = 0 |
| ***Distance to Nearest Outlet*** | C | Distance from Centroid of DA to closest alcohol outlet in metres |
| ***Primary Drinking Restaurant Density (PDRD)*** | R | By Area (# of PDRD in DA/Area of DA), By Population (# PDRD in DA/Population in DA) |
| ***Restaurant Density*** | R | By Area (# of Restaurants in DA/Area of DA), By Population (#of Restaurants in DA/Population in DA) |
| ***Lagged Variables*** | - | Lagged Versions of all Built-Environment Variables were tested in the spatial Durbin model and spatial Durbin error model (Lagged Variable = Built-Environment Variable x Spatial Weight Matrix (W)<br>Lagged Variables measures the built-environment variables in the areas adjacent to the area of interest |
| **B – Binary Variable**<br>**C – Continuous Variable**<br>**R – Ratio Variable** | | |

## 3.2 Exploratory Spatial Data Analysis

### 3.2.1 Global Spatial Autocorrelation and Local Cluster Methods

The first and fourth research question asked: Is there geographical variation in the average total household alcohol expenditure by Dissemination Area in the City of Toronto? If geographical variation is present, where are the hot-spots and cold-spots of expenditure? If there is geographical variation, this indicates that areas with similar levels of alcohol expenditure are clustering together; therefore, spatial regression models should be used instead of OLS regression for examining associations. Furthermore, local clustering methods can be used to identify the cluster locations. In order to answer research question one and four for both licensed premises and purchased in stores expenditures respectively, Global Moran's I measure, a global clustering method, was first used to determine if there was spatial autocorrelation in the expenditure data throughout the entire dataset by rejecting the null hypothesis of no spatial autocorrelation in the expenditure dataset. Spatial autocorrelation is the measure of similarity or association of values over space, in other words, whether areas with similar values for alcohol expenditures will tend to cluster together (22, 110). Positive spatial autocorrelation indicates that clustering of areas with similar values is occurring within the dataset, and as a result, this confirms that there is geographical variation in the variable throughout the dataset. The Global Moran's I measure is calculated with the following formula (23, 110-113):

$$I = \frac{\frac{N}{S_0}\, \Sigma_i \Sigma_j W_{ij} Z_i Z_j}{\Sigma_i Z_i^2}$$

Where $\boldsymbol{W_{ij}}$ is a matrix describing spatial proximity and $\boldsymbol{Z_i Z_j}$ is a matrix describing how similar are values of a particular attribute. $\boldsymbol{W_{ij}}$ is a spatial weight matrix (sometimes called a neighbouring function) which uses binary values to indicate whether a location j is adjacent (considered a neighbour) of the location of interest, location x (111). The definition of what constitutes a neighbour will have a significant effect on this matrix. For studies that use area level data such as the current study, a common definition used is queen contiguity. This specifies that if two areas have a shared boundary line or shared boundary vertex, they are considered neighbours (22). Areas that are considered neighbours are given a value of one whereas areas considered non-neighbours are given a value of zero (114). $\boldsymbol{Z_i Z_j}$ can also be written as $(\boldsymbol{x_i} - \overline{\boldsymbol{x}})(\boldsymbol{x_j} - \overline{\boldsymbol{x}})$, where $\boldsymbol{x_i}$is an attribute in area i and $\overline{\boldsymbol{x}}$ is the mean over all z-values. A value is determined between $\boldsymbol{x_i}$ and all possible surrounding areas ($\boldsymbol{x_j}$) (111). This matrix does not take into consideration whether or not $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$ are considered neighbours.

When $\boldsymbol{W_{ij}}$ and $\boldsymbol{Z_i Z_j}$ are multiplied together the final matrix that is produced accounts for both how similar values are between locations and whether or not the locations are neighbours. As non-neighbours were assigned a zero value, their final value will be zero regardless of the measure of similarity. Neighbouring areas that were assigned a value of one will have their measure of similarity reported, as it will be multiplied by a value of one. After $\boldsymbol{W_{ij}}$ and $\boldsymbol{Z_i Z_j}$are multiplied together, dividing them by the denominator $\boldsymbol{\Sigma_i Z_i^2}$ normalizes the output to a scale with a range of -1 to +1 (115). The other fraction in the equation is $\frac{\boldsymbol{N}}{\boldsymbol{S_0}}$ where N is the total number of observations (in this case the number of areas) in the data set. $\boldsymbol{S_0}$ is exactly the same $\boldsymbol{\Sigma_i \Sigma_j W_{ij}}$, meaning that the total number of observations is divided by the spatial weight matrix. This portion is only included if the spatial weight matrix ($\boldsymbol{W_{ij}}$,) has not been standardized.

The null hypothesis for Global Moran's I is that there is no spatial autocorrelation (113). This is indicated by a negative Global Moran's I ($I$ < 0). A score of -1 indicates perfect negative spatial autocorrelation (116). A classic example of this is a checkerboard pattern where only dissimilar values are next to each other. A Global Moran's I value greater than 0 ($I > 0$) indicates positive spatial autocorrelation, where similar values are clustered together. The closer the value is to +1 the greater the degree of clustering (116).

While the Global Moran's I values are able to quantify the average level of spatial autocorrelation in the data, the global clustering method cannot identify the specific locations of the clusters or the significance of individual clusters as it assumes homogeneity for the entire study region (110). In order to identify specific clusters and their locations, local clustering methods must be applied. Local clustering methods also known as local indicators of spatial association (LISA) (22, 115), are able to consider subsets of the entire dataset to identify specific clusters within the study region. In the current study a LISA method would be able to identify specific DAs that are adjacent to each other with similar values for an attribute. For this study, the local clustering method known as Local Moran's I (LMI) is applied (22, 115). The equation for LMI is:

$$I_i = \frac{Z_i}{m_2} \sum_j W_{ij} Z_j$$

Where $\boldsymbol{Z_i}$ is the standardized value of the variable of interest in subarea *i*,, $\boldsymbol{m_2}$ is the constant for all areas, $\mathbf{Z_j}$ is the standardized value of the variable of interest in neighbouring subareas and $\mathbf{W_{ij}}$ is once again the spatial weight matrix which defines which subareas *j* are considered neighbours to a subarea *i*. As a result, the LMI measures the statistical correlation between the value of a variable in subarea *i* to the value of the same variable in neighbouring subareas. If there is a positive LMI value, this indicates that subareas with similar values of a variable are clustering together (22). This can be either high values clustering together or low values clustering together, which produces hot-spots and cold-spots respectively. A negative LMI value indicates that neighbouring subareas have dissimilar values for the variable of interest (22).

Along with identifying hot-spots and cold-spots, the LMI also provides significance testing by Monte Carlo methods (22). Monte Carlo significance testing is done by generating a null hypothesis reference distribution for the LMI statistic. The generation is done by repeated random assignment of data values. The observed values for the LMI are then compared to the simulated values from the reference distribution. The Monte Carlo *p*-value is then calculated with the formula $\boldsymbol{m(n+1)}$ where *m* is the number of simulated values that exceeded the observed LMI and *n* is the total number of observations (22).

### 3.2.2 Global Spatial Autocorrelation and Local Cluster Analysis Procedure

For both licensed premises and purchased from stores alcohol expenditure datasets, the Global Moran's I statistic was applied to first quantify the average level of spatial autocorrelation in the entire dataset. Both expenditure types showed significant positive spatial autocorrelation. In the next step, local clustering methods were applied to both expenditure types in the form of the Local Moran's I measure, which identified significant hot-spots and cold-spots for each expenditure type. Both the global clustering method and local clustering methods were carried out using GeoDa v1.6.7. The hot-spots and cold-spots were then exported and then imported into ArcGIS 10.3, where high resolution maps were created. Additionally, overlay analysis was carried out by overlaying a map of the major roads and highways in Toronto onto the cluster maps. This helped identify specific streets as boundaries for the clusters. Finally, spatial selection methods were used to identify the recognized Toronto neighbourhoods associated with each hotpot and cold-spot. Figure 3.2.2.1 and Figure 3.2.2.2 provides two flow charts that diagram the entire procedure used for the cluster analysis.

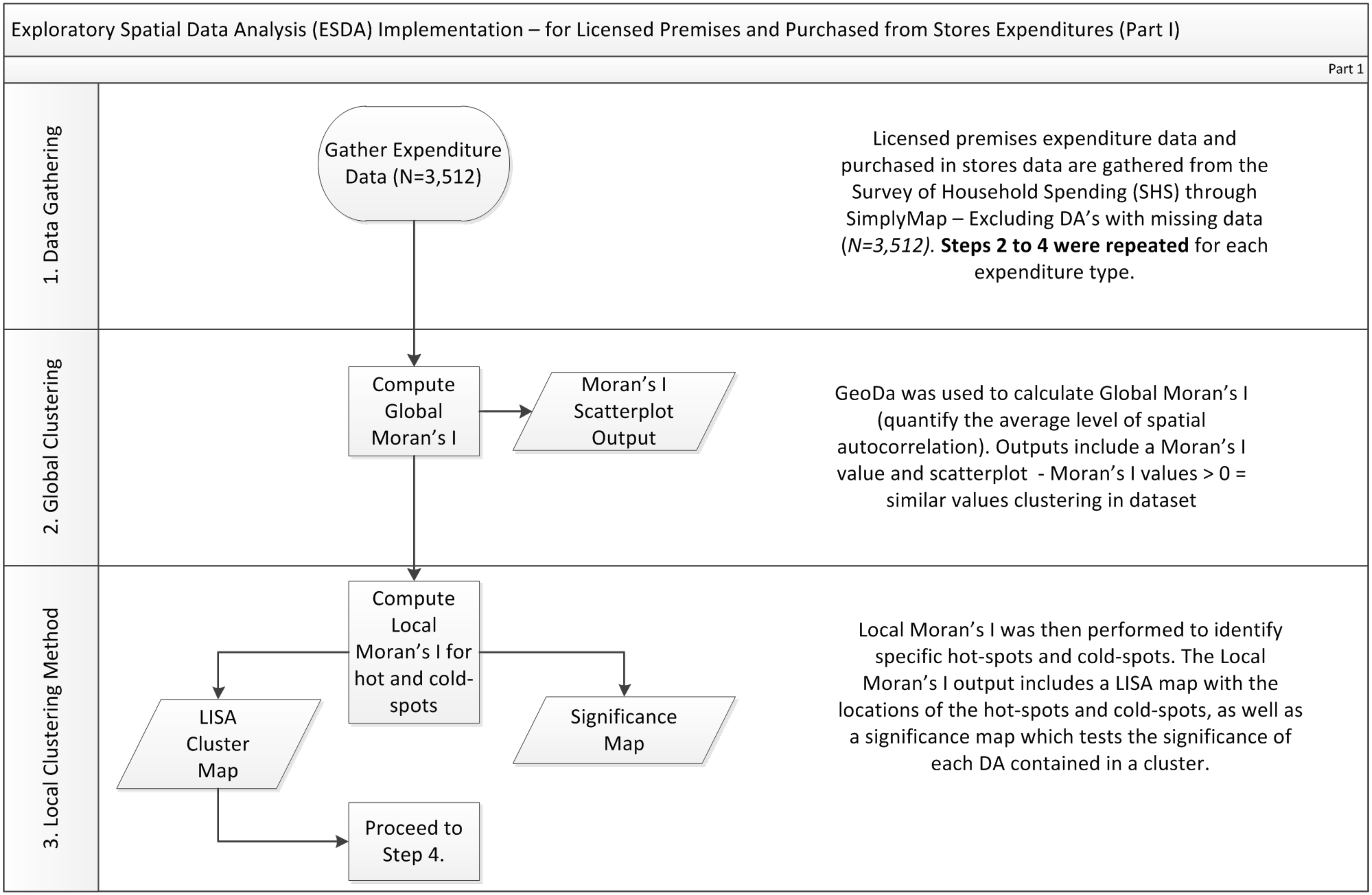



Figure 3.2.2.1: ESDA Implementation - Part 1

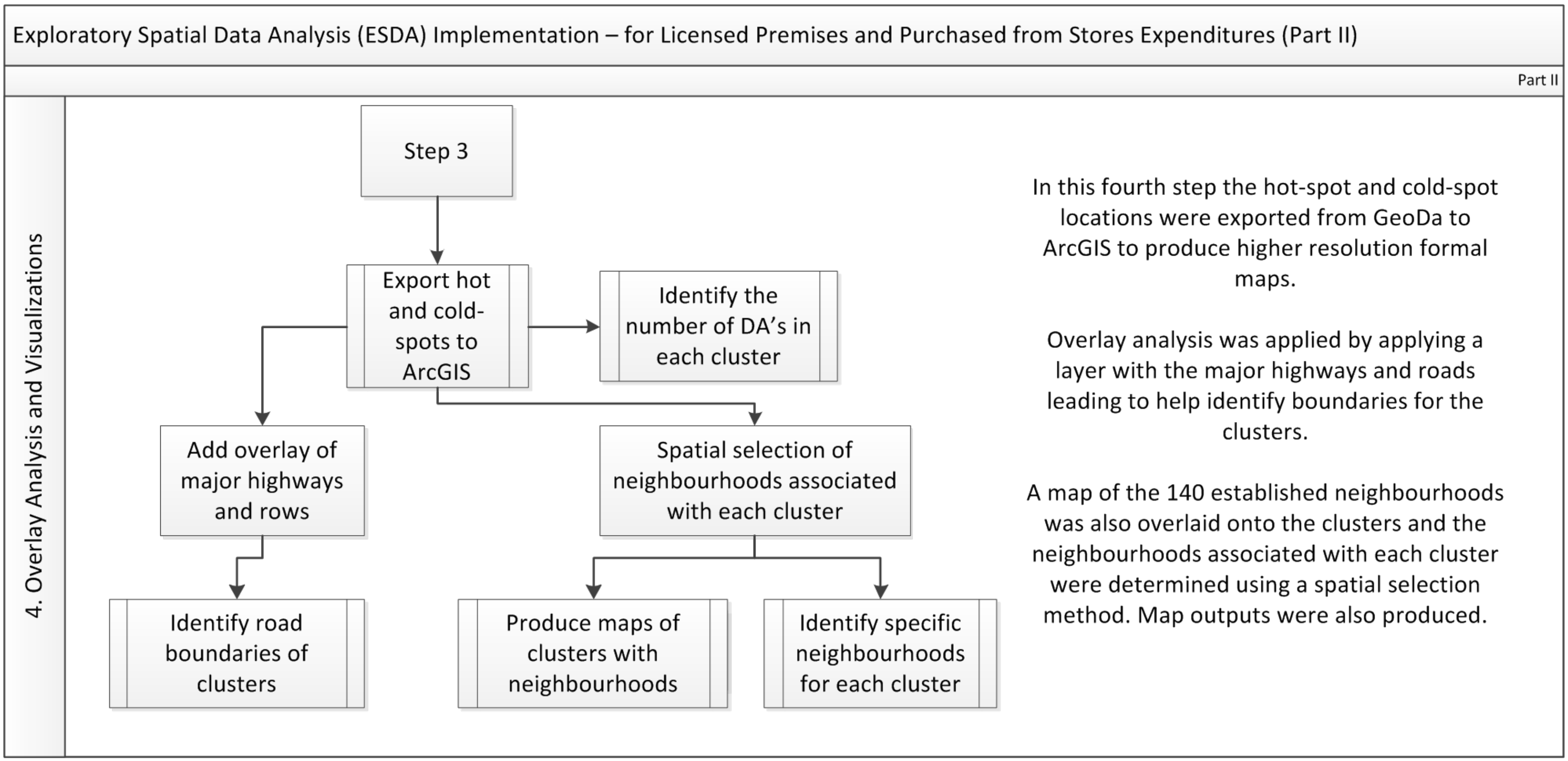



Figure 3.2.2.2: ESDA Implementation - Part 2

## 3.3 Confirmatory Spatial Data Analysis

### 3.3.1 OLS and Spatial Regression Models

Research question two and five asked: What are the socio-economic and built-environment variables that are associated with average total household alcohol expenditure by Dissemination Area in the City of Toronto? Traditionally, the original least squares (OLS) regression model is used to estimate the linear relationship between a dependent variable (outcome variable) and independent variable or multiple independent variables (explanatory variables). OLS regression takes the form:

$$y = B_0 + B_i x_i + \varepsilon$$

Where $\boldsymbol{y}$ is alcohol expenditure, $\boldsymbol{B_0}$ is the egression constant, $\boldsymbol{x_i}$ is an independent variable, $\boldsymbol{B_i}$ is the regression co-efficient or parameter estimate for independent variable $x_i$ and $\varepsilon$ is the error term. Only one independent variable is explored, making this a bivariate OLS regression. An example of this is exploring the relationship between alcohol expenditure and income. This model can be indefinitely expanded to include multiple independent variables, which takes the form:

$$y = B_0 + B_1 x_1 + B_2 x_2 + \cdots + B_n x_n + \varepsilon$$

An example of this is examining the relationship between alcohol expenditure (outcome variable) and income, education and lone-parent families (independent variables).

However, in the presence of positive spatial autocorrelation in the outcome variables, OLS regression is no longer acceptable as it does not control for spatial clustering of the data (23). As a result, OLS parameter estimates can be biased and inefficient, the estimated standard errors can also be biased and the model will be prone to Type I (False Positive) & II (False Negative) errors depending on the type of spatial autocorrelation (87). When a dependent variable in an area is influenced by values of the dependent variable in adjacent areas, this results in what is called a "spillover" effect. Spatial regression models have been developed to address this "spillover" effect and have been widely employed in spatial econometrics especially in examining temporal dependence (117). There are four main spatial regression models: Spatial lag regression model (also known as Spatial autoregressive model or SAR), Spatial error regression model (SEM), Spatial Durbin model (SDM) and Spatial Durbin error model (SDEM) (23, 112, 113, 117-120).

#### *3.3.1.1 Spatial Lag Regression Model (Spatial Autoregressive Model – SAR)*

The spatially lagged $y$ model or spatial lag regression model (also called "spatial autogregressive" model by Anselin (121)) is used when values of the dependent variable ($\boldsymbol{y}$) for a unit ($\boldsymbol{i}$) are thought to be directly influenced by the values of $\boldsymbol{y}$ in neighbouring units of $\boldsymbol{i}$ (23, 122). In this study, if the value of average total alcohol expenditure for a Dissemination Area $\boldsymbol{i}$ is considered to be influenced by values of average total alcohol expenditure in neighbouring Dissemination Areas $\boldsymbol{j}$, then the spatial lag regression would be an appropriate spatial model to apply. In order to account for this effect, a spatially lagged dependent variable ($\boldsymbol{\rho W y}$) is included on the right-hand side of the equation as an explanatory variable. For $\boldsymbol{\rho W y}$ , $\boldsymbol{\rho}$ is a scalar parameter indicating the strength of the spatial association (also called spatial autoregressive parameter) (119), $\boldsymbol{W}$ is the connectivity matrix or spatial weight matrix that defines what areas are neighbours for a specific area $\boldsymbol{I}$ ***(23)***. The full form of the equation is:

$$Y = \rho W y + X\beta + \epsilon$$

This model provides additional information about the spatial structure of alcohol expenditure as the model is indicating that expenditure values in one area are influenced by the expenditure values in neighbouring areas as well (123).

*3.3.1.2 Spatial Error Regression Model (SEM)*

A different way to incorporate spatial dependence into the OLS regression model is to include it into the error term. The spatial error regression model incorporates a spatially lagged error term as part of the explanatory variables. The spatially lagged error term calculates the mean regression error in neighbouring Dissemination Areas. The notation for the spatially lagged error term is $\boldsymbol{\lambda W \xi}$, where $\boldsymbol{\lambda}$ is the spatial autoregressive parameter that quantifies the strength of the spatial association, $\boldsymbol{W}$ is the connectivity matrix or spatial weight matrix which defines what areas are neighbours for a specific area $\boldsymbol{i}$ and $\boldsymbol{\xi}$ is the spatial component of the error term (23). In this model $\boldsymbol{\epsilon}$ is the error term for the regression model. The model is described by:

$$y = X\beta + \lambda W\xi + \epsilon$$

The spatial error regression provides a different interpretation of spatial dependence as the model assumes that the spatial correlation ($\lambda$) within the data is not useful and primarily a nuisance, as it does not add additional information about spatial patterning in the model (23). However, although the spatial error model does not offer direct information regarding the spatial patterning, the significance of the spatial error term offers diagnostic information, as correlation between the errors of a model can be purely coincidental or may indicate a misspecification in the systematic part of the model ($X\beta$), namely a significant independent variable is missing (23, 123).

### *3.3.1.3 Spatial Durbin Model (SDM)*

An extension to the spatial lag regression model is the spatial Durbin model (SDM). In the spatial Durbin model there is a spatially lagged dependent variable added, however, additional lagged independent variables are added as well ($\boldsymbol{W_1 X \beta_2}$). The formula is given as (112, 113, 117, 119, 120):

$$y = \rho W_1 y + \alpha\iota_n + X\beta_1 + W_1 X\beta_2 + \varepsilon$$

Using this model in the context of this study, the dependent variable would be alcohol expenditure, an independent variable would be alcohol outlet density and a lagged independent variable would be lagged alcohol outlet density or the mean alcohol outlet density in neighbouring areas.

### *3.3.1.4 Spatial Durbin Error Model (SDEM)*

A fourth and final common model is the spatial Durbin error model. Just as the spatial Durbin model is an extension of the spatial lag model (SAR), the spatial Durbin error model is an extension of the spatial error model. The spatial Durbin error model is given by the formula (117, 119, 120):

$$y = X\beta_1 + WXB_2 + u$$

$$u = \lambda Wu + \varepsilon$$

The spatial Durbin error model includes both a spatially lagged error term ($\boldsymbol{\lambda W u}$)and spatially lagged independent variables ($\boldsymbol{WXB_2}$).

### 3.3.2 Multivariate Spatial Regression Procedure

To answer research question two and five, multivariate spatial regressions were carried out to identify socio-economic and built-environment determinants for both alcohol expenditure types. The multivariate spatial regression procedure began with descriptive statistics, and correlation analysis was also performed to describe the distribution of the dataset and to deal with potential issues of multicollinearity. Preliminary bivariate OLS Regressions all displayed positive spatial autocorrelation; therefore, bivariate regressions were calculated using the spatial error model. Significant explanatory variables were then placed in the same model to create a multivariate spatial error model. The number of insignificant explanatory variables was reduced with backwards stepwise regression (**$\alpha = 0.10$**). In order to carry out the spatial regressions, a spatial weight matrix file was created to model the spatial structure of the dataset. The spatial weight matrix featured a row-standardized first-order queen contiguity definition to define neighbours for each Dissemination Area. GeoDa v1.6.7 was used to create the spatial weight matrix file. All OLS and spatial regressions were completed in GeoDa v1.6.7 before being re-done in the statistical software R using a variety of packages. The regressions were performed again in R to add the spatial Durbin model and spatial Durbin error models and for increased graphics capabilities. The four different spatial regression models were tested and the best model was refined and determined by the highest log-likelihood value. Figure 3.3.4.1 and Figure 3.3.4.2 provide two flowcharts detailing the regression steps.

### 3.3.3 Log Transformed Outcome Variable

As the outcome expenditure variables in this study have been log transformed, the direct interpretation of the parameter estimates is the change in the log transformed alcohol expenditure values from one-unit change in the explanatory variable, while controlling for all other variables.

However, this direct interpretation is difficult to understand in terms of the actual change in alcohol expenditure values due to the log transformation. To find the effect of one-unit change in an explanatory variable on the original alcohol expenditure outcome variable, the regression coefficients need to be modified using the following formula (124):

$$\%\Delta y = 100 \times \left(e^{\beta x} - 1\right)$$

Where x is number of unit changes for the explanatory variable, beta is the parameter estimate and y refers to changes to the geometric mean of the original alcohol expenditure value. The new interpretation is: For a one-unit change in $X$ we expect the outcome variable $Y$ (alcohol expenditure) to change by $100 \times \left(e^{\beta x} - 1\right)$ percent.

### 3.3.4 Examining for Remaining Spatial Autocorrelation

Research question three and six asked: Is there remaining spatial autocorrelation in the licensed premises and purchased in stores alcohol expenditures after controlling for small-area socio-economic and built-environment factors? After modeling the spatial variables, any remaining geographic variation in the residuals of the final spatial models can be quantified. In order to do so to answer research question three and six, the Moran's I statistic was carried out on the residuals of best spatial model identified for each expenditure type. The results showed that spatial autocorrelation had been accounted for in both models and as a result, was no longer an issue.

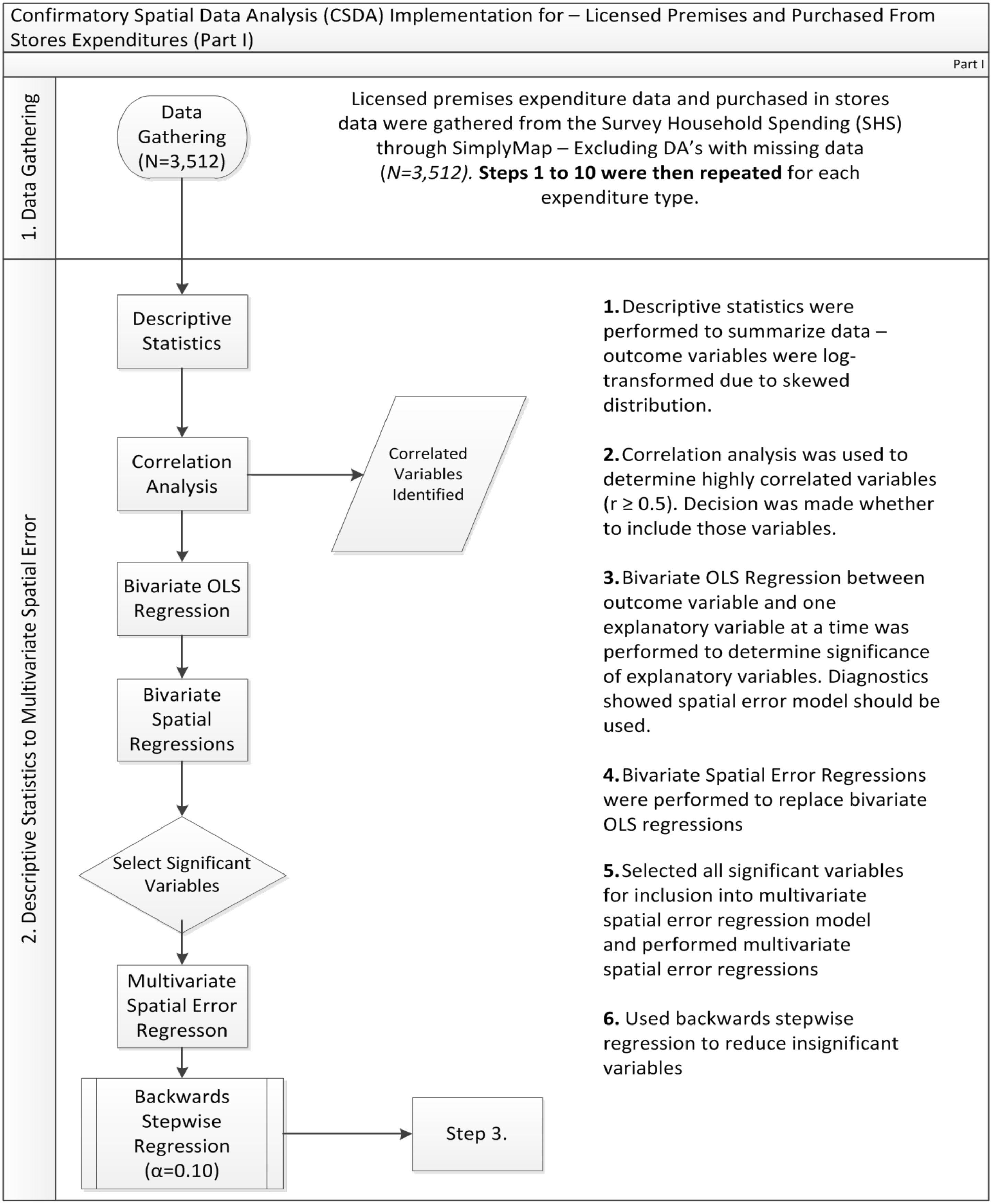



Figure 3.3.4.1: CSDA Implementation - Part 1

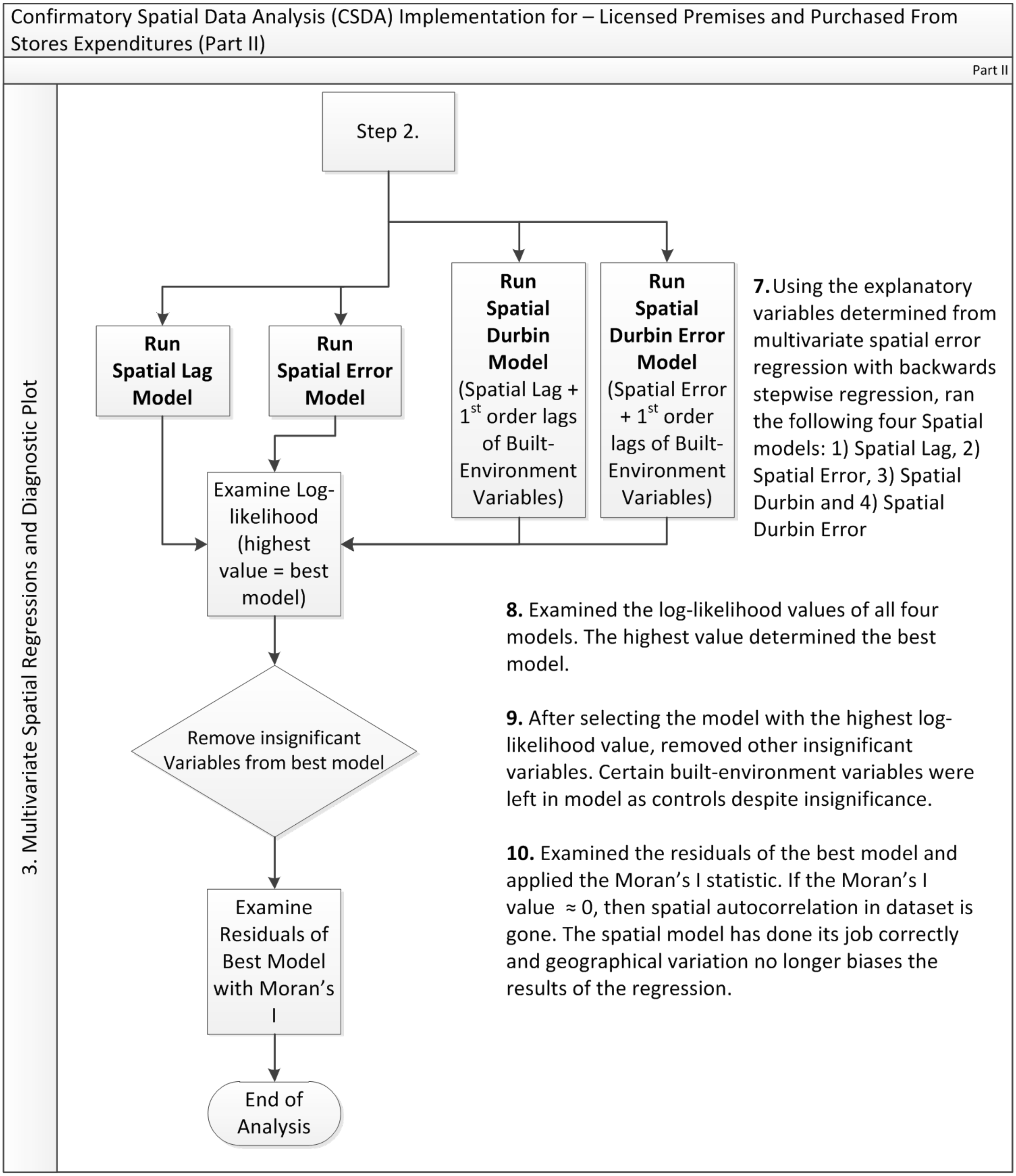



Figure 3.3.4.2: CSDA Implementation - Part 2

# Chapter 4: Results

## 4.1 Descriptive statistics

In 2010, the minimum average household alcohol expenditure in licensed premises was $47.54 and the maximum was $2963.02. The average expenditure per household was $337.51. For average household alcohol expenditure from purchases in stores, the minimum was $81.18 and the maximum was $5894.13. The average expenditure per household was $629.80. As both of these variables were skewed to the right (see Appendix A1 for histograms), log transformed versions of each expenditure variable were created for the regression analysis. Table 3.3.4 1 provides descriptive statistics for the expenditure outcome variables, the socio-economic independent variables and the built-environment independent variables.

### 4.1.1 Descriptive Statistics Summary Table

Table 4.1.1-1: Table of descriptive statistics for all outcome and explanatory variables

| Variable | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| Expend. in Licensed Premises | 3512 | 47.54 | 2963.02 | 337.51 | 291.01 |
| Expend. From Purchases in Store | 3512 | 81.18 | 5894.13 | 629.80 | 493.49 |
| Log Transformed - LP | 3512 | 3.86 | 7.99 | 5.56 | 0.70 |
| Log Transformed - PFS | 3512 | 4.40 | 8.68 | 6.25 | 0.60 |
| % 20 to 44 years | 3512 | 0.04 | 0.83 | 0.38 | 0.09 |
| Male to Female Ratio | 3512 | 0.00 | 33.68 | 0.96 | 0.57 |
| % White | 3512 | 0.00 | 100.00 | 59.72 | 26.62 |
| % Ethnicity Black | 3512 | 0.00 | 18.93 | 0.37 | 1.32 |

| Variable | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| % Ethnicity Chinese CHN | 3512 | 0.00 | 94.62 | 10.22 | 15.76 |
| % Ethnicity Filipino FLP | 3512 | 0.00 | 60.86 | 3.11 | 5.38 |
| % Ethnicity South Asian SA | 3512 | 0.00 | 30.14 | 0.42 | 1.59 |
| % Blue-collar | 3512 | 0.00 | 0.69 | 0.17 | 0.12 |
| % White-collar | 3512 | 0.07 | 0.93 | 0.53 | 0.15 |
| Median Household Income - BT | 3512 | 11807.00 | 551293.00 | 66419.95 | 35302.93 |
| Median Household Income - AT | 3512 | 11776.00 | 335039.00 | 56303.79 | 25201.81 |
| % Transfer Payments | 3512 | 0.00 | 56.30 | 11.17 | 7.37 |
| % Apartments – 5 or more floors | 3512 | 0.00 | 100.00 | 19.10 | 33.69 |
| % Apartments – Less than 5 floors | 3512 | 0.00 | 100.00 | 19.04 | 23.95 |
| % Apartments - Duplex | 3512 | 0.00 | 61.39 | 6.60 | 7.92 |
| % Row Houses | 3512 | 0.00 | 100.00 | 5.62 | 14.00 |
| % Semi-Detached | 3512 | 0.00 | 100.00 | 9.62 | 16.41 |
| % Single Detached | 3512 | 0.00 | 100.00 | 39.87 | 34.50 |
| % Dwellings Rented | 3512 | 0.00 | 100.00 | 35.33 | 29.96 |
| % Married Couples | 3512 | 0.00 | 100.00 | 61.91 | 16.53 |
| % Lone-Parent | 3512 | 0.00 | 50.00 | 14.93 | 7.87 |
| % Post-Secondary Education | 3512 | 0.00 | 0.94 | 0.48 | 0.16 |
| Avg. Number of Person in Private Dwelling | 3512 | 1.11 | 6.46 | 2.69 | 0.59 |
| % Aboriginal Identity | 3512 | 0.00 | 50.00 | 0.64 | 1.73 |

| Variable | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| % Residential Land | 3512 | 0.00 | 1.00 | 0.73 | 0.32 |
| Lagged % Residential Land | 3512 | 0.00 | 1.00 | 0.69 | 0.18 |
| % Commercial Land | 3512 | 0.00 | 1.00 | 0.02 | 0.11 |
| Lagged % Commercial Land | 3512 | 0.00 | 0.58 | 0.02 | 0.05 |
| Highway Intercept | 3512 | 0.00 | 1.00 | 0.05 | 0.22 |
| Lagged Highway Intercept | 3512 | 0.00 | 1.00 | 0.07 | 0.15 |
| Subway Intercept | 3512 | 0.00 | 1.00 | 0.06 | 0.23 |
| Lagged Subway Intercept | 3512 | 0.00 | 1.00 | 0.07 | 0.16 |
| Outlet Density by Population | 3512 | 0.00 | 10.00 | 0.01 | 0.19 |
| Lagged Outlet Density by Population | 3512 | 0.00 | 129536.88 | 7524.22 | 5867.16 |
| Outlet Density by Area | 3512 | 0.00 | 97.16 | 0.52 | 3.94 |
| Lagged Outlet Density by Area | 3512 | 0.00 | 26.43 | 0.56 | 1.80 |
| Primary Drinking Restaurant Density By Area | 3512 | 0.00 | 116.78 | 1.10 | 6.46 |
| Lagged Primary Drinking Restaurant Density By Area | 3512 | 0.00 | 38.79 | 1.13 | 3.43 |
| Eating Restaurant Density by Area | 3512 | 0.00 | 826.78 | 21.54 | 59.30 |
| Lagged Eating Restaurant Density by Area | 3512 | 0.00 | 493.35 | 21.53 | 41.11 |
| Pop Density by Area KMS | 3512 | 9.13 | 423688.90 | 8491.02 | 12303.91 |
| Distance to Nearest Outlet | 3512 | 3.96 | 3121.90 | 921.02 | 531.49 |

### 4.1.2 Bivariate Analysis

A correlation matrix with all of the dependent and explanatory variables was also created to investigate for multicollinearity. Explanatory variables that had a Pearson coefficient of greater that 0.5 (r>0.5) were considered to be highly correlated. Table 4.1.1 1 presents the variable pairs that were found to be correlated:

Table 4.1.1-1 - Highly Correlated Variables

| | Correlated Pair | Pearson's R |
|---|---|---|
| 1. | Median Before-Tax Household Income and Median After-Tax Household Income | 0.974 |
| 2. | Median Before-Tax Household Income and % of White-collar | 0.512 |
| 3. | Median After-Tax Household Income and % of White-collar | 0.502 |
| 4. | % Transfer Payment and Percentage Blue-collar | 0.530 |
| 5. | % Single Detached Housing and Median Before-Tax Household Income | 0.548 |
| 6. | % Single Detached Housing and Median After-Tax Household Income | 0.584 |
| 7. | % Rented Dwellings and Percentage Apartments – 5 floors or more | 0.627 |
| 8. | % Post-Secondary and % White-collar | 0.724 |
| 9. | Average Number of Persons in Private Households and Percentage 0 to 19 years | 0.615 |
| 10. | % Married Couples and % Single Detached Homes | 0.503 |
| 11. | % Blue-collar and % White-collar | -0.770 |
| 12. | % Transfer Payment and % White-collar | -0.607 |
| 13. | % Transfer Payment and Median After-Tax Household | -0.514 |
| 14. | % Single-Detached and % 20 to 44 years old | |
| 15. | % Single-Detached and % Percentage Apartments – 5 floors or more | -0.547 |
| 16. | % Rented Dwellings and Median After-Tax Income | -0.585 |
| 17. | % Post-Secondary Education and % Blue-collar | -0.715 |
| 18. | % Post-Secondary and % Transfer Payment | -0.654 |

Pairs two to ten and twelve to sixteen were variables were highly correlated variables with values ranging from $|r| = 0.502$ to $|r| = 0.724$ , however, since they measured different dimensions (i.e. % single detached housing and median before-tax household income looks at housing and income respectively), these variables were all included in the next stage of OLS regression analysis. Pair one and eleven presented a different case as the two income variables ($r = 0.974$) and the two occupation variables ($r = -0.770$) were highly correlated variables in the same domain. For pair one, median after-tax income was chosen over before-tax income as it was more highly correlated to the alcohol expenditure variables. As well, after-tax income was expected to be more reflective of actual disposable income for households. For pair eleven, % white-collar occupations was selected over % blue-collar occupations for inclusion in the spatial models, as white-collar occupations showed a higher correlation to the alcohol expenditure variables.

## 4.2 Licensed Premises Expenditure Data

### 4.2.1 Global Spatial Autocorrelation

To answer research question one, Global Moran's I was first calculated for the licensed premises expenditure data to determine if geographic variation was present in the dataset. For licensed premises alcohol expenditure, the Global Moran's I value was 0.633957 and was significant at the $p = 0.001$ significance level (999 Permutations). As positive spatial autocorrelation values range from > 0 to +1, this value represented a significant and high degree of positive spatial autocorrelation. This result confirmed that there was significant geographic variation in the alcohol expenditure dataset by DA in the City of Toronto. Figure 4.2.1.1 presents the Moran's I scatterplot used to calculate the Global Moran's I value for the licensed premises alcohol expenditure data. The licensed premises expenditure values (x-axis) were plotted against expenditure values in nearby areas (y-axis), and the resulting positive slope indicated the presence of geographical variation.

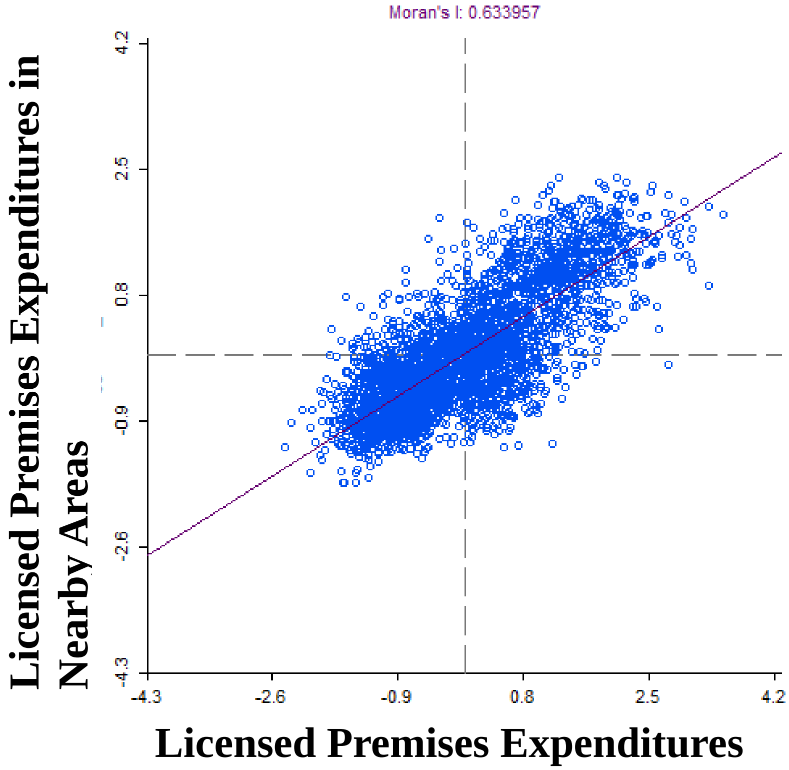



Figure 4.2.1.1: Moran's I Scatterplot for Licensed Premises Alcohol Expenditures: The Moran's I statistic plots the licensed premises expenditure values (x-axis) against expenditure values in nearby areas (y-axis). The resulting correlation value is the Moran's I value. Monte Carlo testing using 999 permutations was used to produce a p-value that determined whether the points plotted form a significant cluster. The Moran's I values were significant at the 0.001 significance level.

### 4.2.2 Local Spatial Autocorrelation – Hot-Spots and Cold-Spots

Having determined that there was geographic variation in the alcohol expenditure data, the next step was to answer the second half of research question one, which asked: Where were the locations of the hot-spots and cold-spots of alcohol expenditure? Figure 4.2.2.1 shows a significance map highlighting that six different clusters have been identified that are significant at the $p = 0.01$ significance level. Figure 4.2.2.2 identifies the clusters by hot-spot (red) and cold-spot (blue). There were a total of three hot-spots and three cold-spots identified for licensed premises alcohol expenditures by using the Local Moran's I statistic.

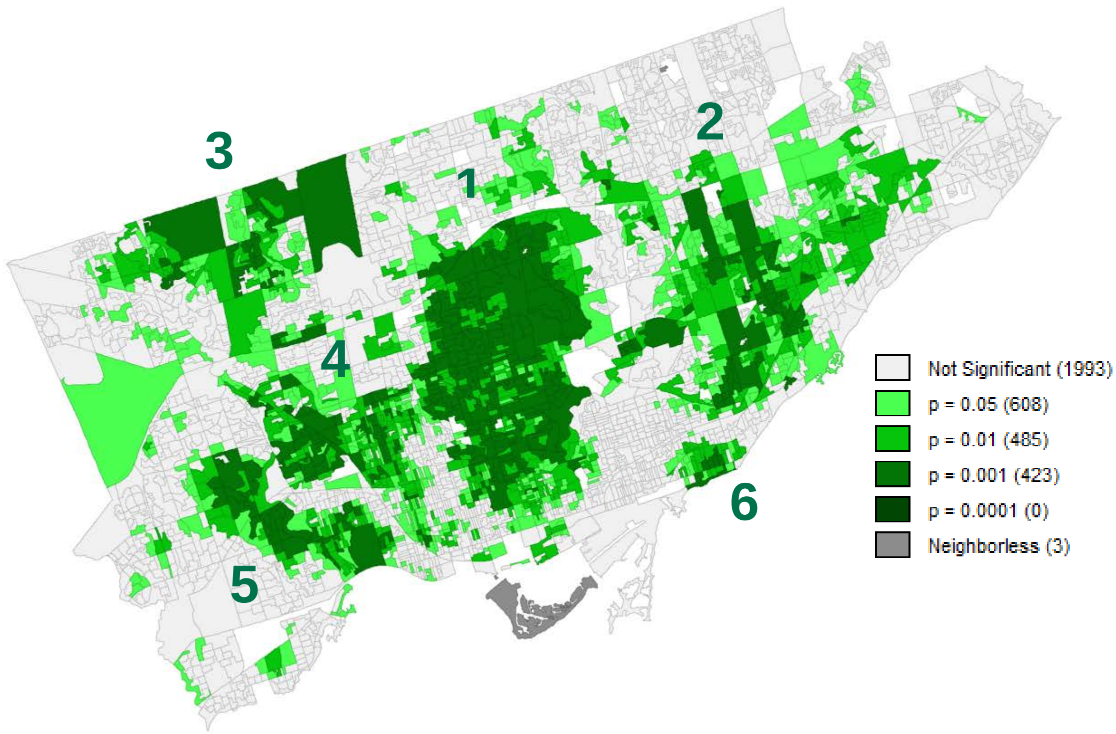



Figure 4.2.2.1: Significance Map for Clusters Identified for Licensed Premises Alcohol Expenditures: This map shows the significance of the six clusters identified for licensed premises alcohol expenditure. The darker the green colour, the more significant the DA is as part of the cluster.

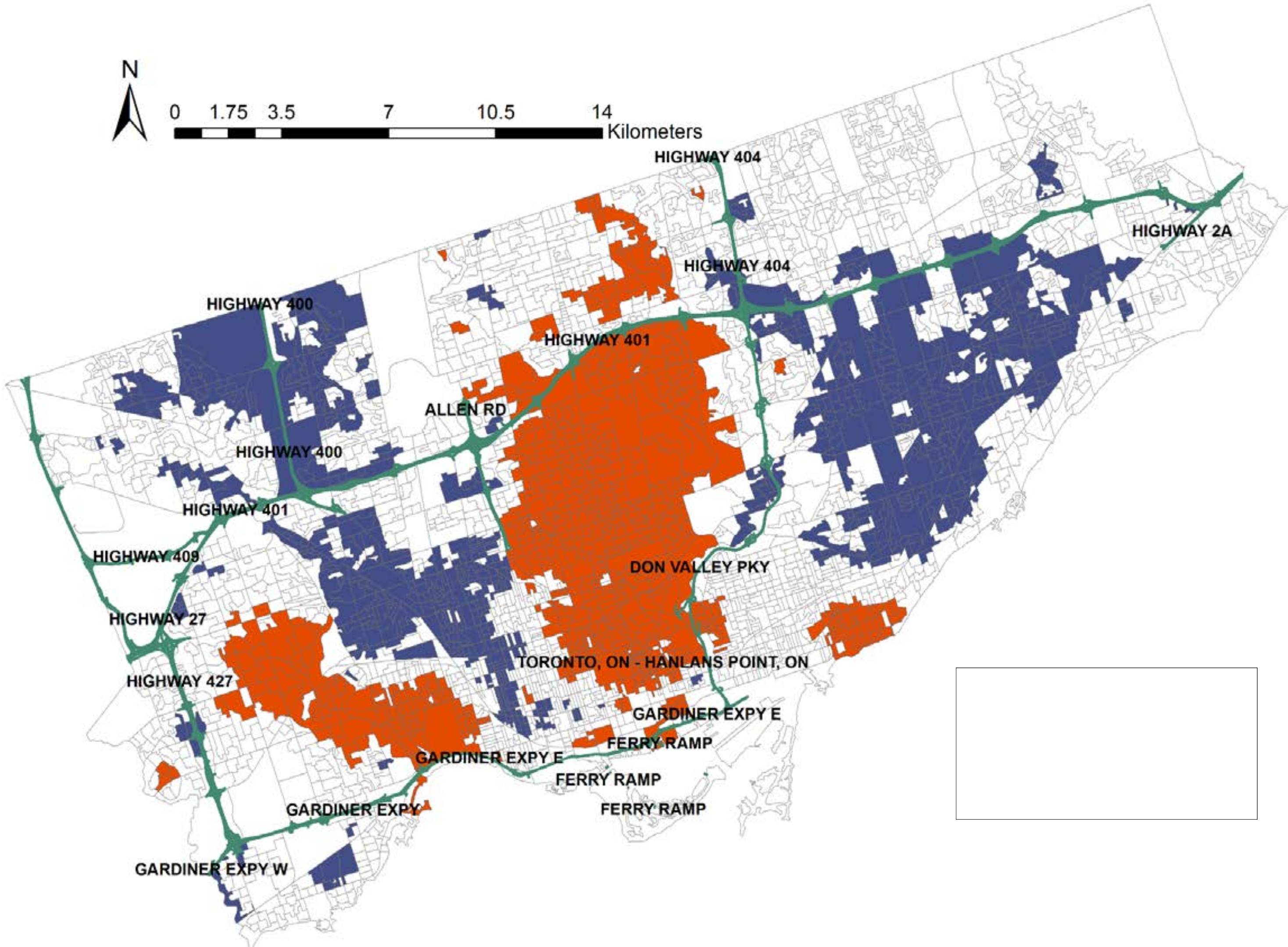



Figure 4.2.2.2: Hot-Spot and Cold-Spot Map for Licensed Premises Alcohol Expenditures: This map shows the hot-spots (areas with high levels of alcohol expenditure clustered together) in red and the cold-spots (areas with low levels of alcohol expenditure clustered together) in blue. The major highways have also been overlaid to provide a sense of location for the clusters.

### 4.2.3 Changing Scales – Dissemination Area to Municipal Neighbourhoods

To extend the results of the analysis to policy-relevant neighbourhood area units, spatial selection methods were used to link the Dissemination Areas in hot-spots and cold-spots to corresponding municipal neighbourhoods. Figure 4.2.3.1 shows the municipal neighbourhoods corresponding to the hot-spot Dissemination Areas and Table 4.2.3-1 provides summary information for the hot-spot clusters. Figure 4.2.3.2 shows the municipal neighbourhood corresponding to the cold-spot Dissemination Areas and Table 4.2.3-2 provides summary information for the cold-spots.

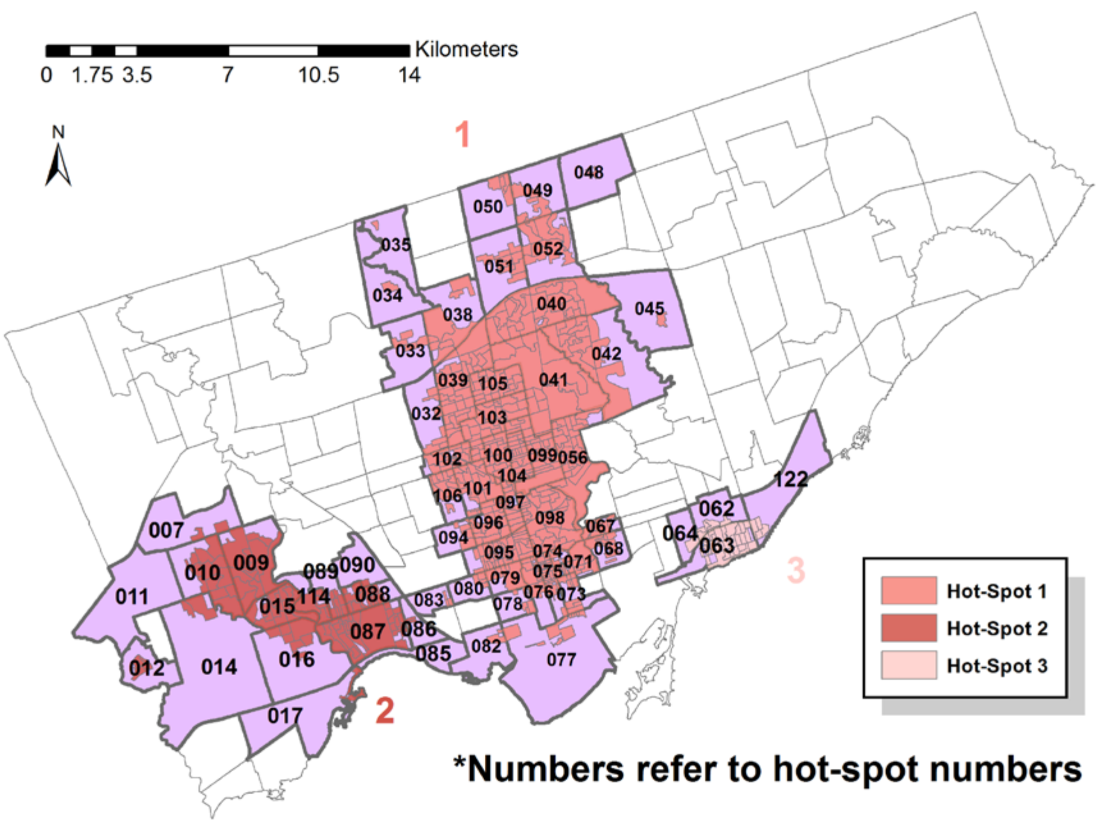



Figure 4.2.3.1 – Hot-Spots with Associated Neighbourhoods: The DAs associated with hot-spots (in shades of red colour) were specifically extracted and overlaid with a map of the 140 municipal neighbourhoods. The neighbourhoods that are associated with these Dissemination Areas are highlighted in pink and are labeled with their neighbourhood number (in black).

Table 4.2.3-1 - Licensed Premises Expenditures Hot-Spots - Summary

| Cluster # | Type | North Boundary | East Boundary | South Boundary | West Boundary | Number of DA | Number of Neighbourhoods |
|---|---|---|---|---|---|---|---|
| **1** | Hot | Steeles Ave | Victoria Park Ave. | Lake Ontario | Dufferin Street/Allen Road/ | 544 | 42 |
| **2** | Hot | The Westway | Landsdowne Ave | Lake Ontario | Etobicoke Creek | 112 | 16 |
| **3** | Hot | Gerrard Street East | Warden Avenue | Lake Ontario | Coxwell Avenue | 38 | 4 |

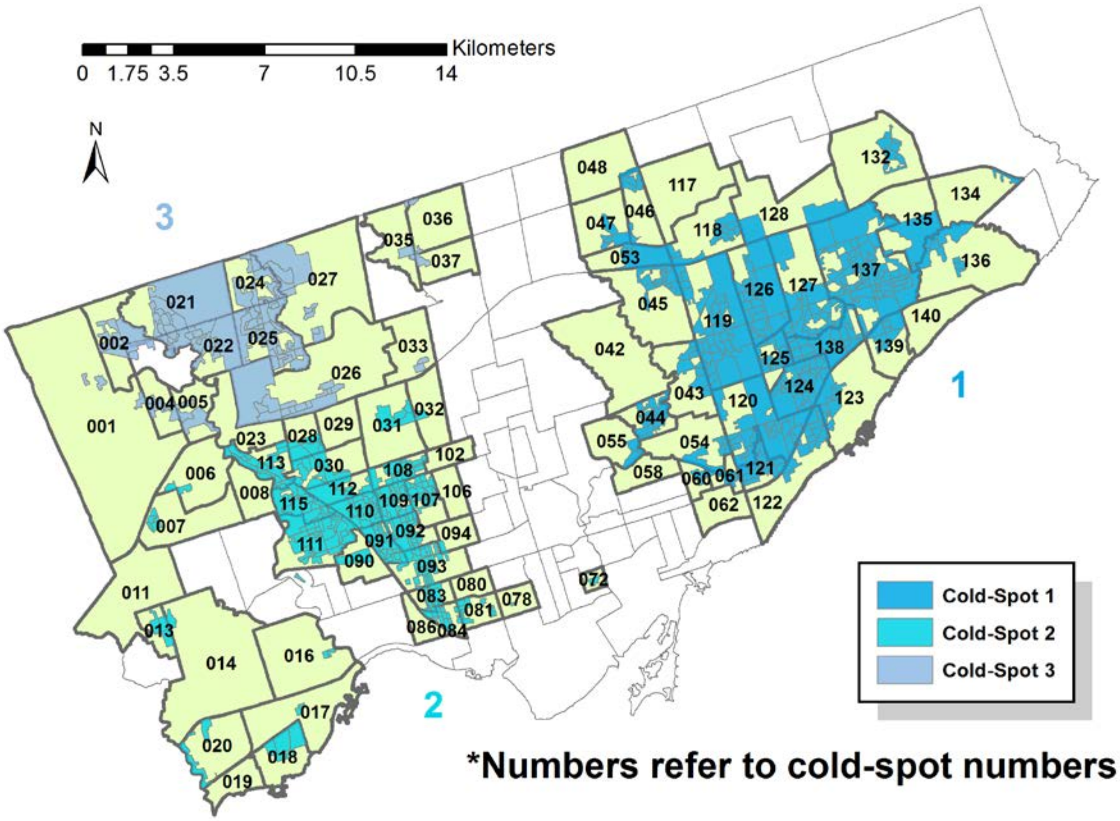



Figure 4.2.3.2 - Cold-Spots and Associated Neighbourhoods: The DAs associated with cold-spots (in teal and blue colours) were specifically extracted and overlaid with a map of the 140 municipal neighbourhoods. The neighbourhoods that are associated with these Dissemination Areas are highlighted in green and are labeled with their neighbourhood number (in black).

Table 4.2.3-2 - Licensed Premises Expenditures Cold-Spots - Summary

| Cluster # | Type | North Boundary | East Boundary | South Boundary | West Boundary | Number of DA | Number of Neighbourhoods |
|---|---|---|---|---|---|---|---|
| **1** | Cold | Steeles Ave | Port Union Rd. | Kingston Road | Leslie Street/Laird Drive | 328 | 34 |
| **2** | Cold | Highway 401 | Allen Road | Lakeshore Blvd W. | Renforth Dr./The West Mall | 276 | 39 |
| **3** | Cold | Steeles Ave | Yonge Street | Highway 401 | Highway 27 | 120 | 14 |

### 4.2.4 Spatial Regressions

The next step after determining the hot-spot and cold-spot locations was to examine the association between alcohol expenditure and the explanatory variables at the DA level in order to answer research question two. Specifically, research question two asked: What were the small-area level socio-economic and built-environment variables associated with alcohol expenditure? As the Global Moran's I measure indicated a significant amount of spatial autocorrelation, multivariate OLS regression models were no longer appropriate for the analysis. Four different spatial regressions models were used on the expenditure data and were then compared (Table 4.2.4-1). The best fitting model to answer research question two for licensed premises expenditure was the spatial Durbin error regression model as it had the largest log-likelihood value, indicating the best fit (Table 4.2.4-1). The regression coefficients and significance of coefficients are provided in the table. The bivariate spatial error regressions can be found in Appendix A2.

Table 4.2.4-1: Spatial Regressions for Licensed Premises Expenditures

**Spatial Regressions for Licensed Premises Expenditures**

| | ***Dependent variable:*** | | | |
|---|---|---|---|---|
| | **Licensed Premises Expenditure** | | | |
| | ***spatial autoregressive*** **(1)** | ***spatial error*** **(2)** | ***Spatial Durbin*** **(3)** | ***Spatial Durbin error*** **(4)** |
| **Socio-Economic Variables** | | | | |
| Percentage White | 0.002*** | 0.004*** | 0.002*** | 0.004*** |
| Percentage of Filipino Ethnicity | -0.005*** | -0.003* | -0.005*** | -0.003* |
| Percentage of Persons in White-collar Occupations | 0.387*** | 0.566*** | 0.389*** | 0.565*** |
| Median Household Income - After-Tax | 0.00001*** | 0.00001*** | 0.00001*** | 0.00001*** |
| Percentage of Income from Transfer Payments | -0.020*** | -0.018*** | -0.019*** | -0.018*** |
| Percentage of Apartments - Less than 5 floors | 0.001*** | 0.001*** | 0.001*** | 0.001*** |
| Percentage of Single Detached Houses | 0.001** | 0.002*** | 0.001*** | 0.003*** |
| Percentage of Row Houses | 0.004*** | 0.005*** | 0.004*** | 0.005*** |
| Percentage Post-Secondary Education | 0.336*** | 0.313*** | 0.297*** | 0.290*** |
| Average Number of Persons in Private Households | -0.043* | -0.044* | -0.025 | -0.034 |
| **Built-Environment Variables** | | | | |
| Subway Intercepts | 0.044 | 0.051 | 0.007 | 0.047 |
| Lagged Subway Intercepts | - | - | 0.078 | 0.244*** |
| Primary Drinking Restaurant Density by Area (KM - Squared) | 0.0004 | -0.0004 | -0.001 | 0.0005 |
| Lagged Primary Drinking Restaurant Density by Area (KM - Squared) | - | - | 0.004 | 0.007* |
| Restaurant Density (KM - Squared) | 0.001*** | 0.0004** | 0.00004 | 0.0003* |
| Lagged Restaurant Density (KM - Squared) | - | - | 0.001*** | 0.001** |
| Constant | 2.699*** | 4.669*** | 2.640*** | 4.606*** |
| **Spatial Autoregressive Term** | | | | |
| Rho | 0.412*** | - | 0.409*** | - |
| Lambda | | 0.658*** | | 0.637*** |
| **Log Likelihood** | -1,656.975 | -1,599.265 | -1,629.309 | -1,578.062 |
| *Note:* | | | | *p<0.05; **p<0.01; ***p<0.001 |

### 4.2.5 Final Diagnostics – Testing for Remaining Spatial Dependence

Research question three asked: Is there any remaining geographical variation in the expenditure after controlling for socio-economic and built-environment factors? Therefore, to answer research question three, the final step of the analysis was to determine whether there was any remaining spatial autocorrelation in the model or if the spatial regression model had accounted for it all. The residuals for the licensed premises spatial Durbin error model were tested with the Global Moran's I statistic and produced a value of -0.045. This represents a very significant decrease from the Global Moran's I value for the linear model residuals (0.290). Additionally, as the value was very close to zero, this indicated that geographic variation in the data no longer presented an issue and proved that the spatial regression model has performed as expected by removing the geographical variation in the data. Figure 4.2.5.1 presents the Moran's I scatterplot of the residuals.

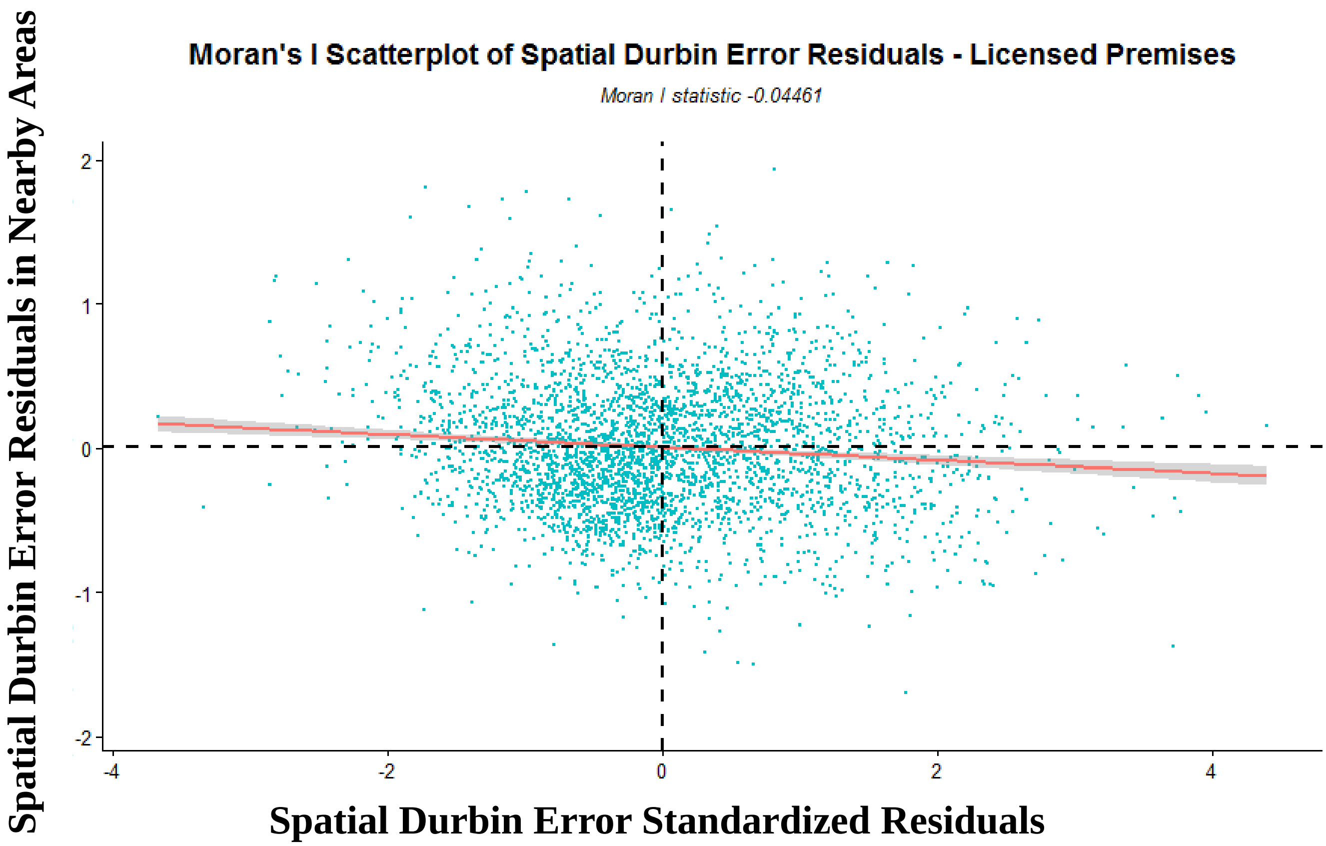



Figure 4.2.5.1: Moran's I Scatterplot of SDEM Residuals for Licensed Premises Alcohol Expenditures: The standardized residuals for the SDEM model (x-axis) are plotted against the standardized residuals in neighbouring areas for the SDEM model. The resulting correlation is the Moran's I value. The value of -0.045 indicates that geographical variation is no longer affecting the model as it is very close to zero.

### 4.2.6 Interpreting the Coefficients for the Spatial Durbin Error Model – Licensed Premises Alcohol Expenditures

#### *4.2.6.1 Socio-Economic Variables*

For ethnicity measures, percentage White and percentage of Filipino ethnicity were the two significant variables in the final model. For percentage White, a one-unit increase in percentage White is expected to increase alcohol expenditures by 0.44%. For percentage of Filipino ethnicity, a one-unit increase in percentage of Filipino ethnicity is expected to decrease alcohol expenditures by 0.28%. For occupation, percentage of white-collar workers was significantly associated with alcohol expenditures. A one-unit increase in percentage of white-collar workers is expected to increase expenditures by 78.57%. Median after-tax incomes was significantly positively associated with alcohol expenditures, while conversely, social assistance measured as percentage of income from transfer payments was negatively associated with alcohol expenditures. For median after-tax income, a one-unit increase in median after-tax income is expected to increase expenditures by 0.0006%. For transfer payments, a one-unit increase in percentage of income from Transfer Payments is expected to decrease expenditures by 1.74%. For housing measures, percentage of apartments with less than five floors, percentage of single-detached houses and percentage of row houses were all positively associated with alcohol expenditure.  For apartments with five floors or less, a one-unit increase in the percentage of apartments with less than five floors is expected to increase expenditures by 0.13%. For single-detached homes, a one-unit increase in the percentage of single-detached houses is expected to increase expenditures by 0.23%. For row houses, a one-unit increase in the percentage of row houses is expected to increase expenditures by 0.48%. Lastly, post-secondary education was also significantly positively associated with alcohol expenditures, as a one-unit increase in post-secondary education is expected to increase expenditures by 35.00%.

#### *4.2.6.2 Built-Environment Variables*

For subways, both subways and lagged subways (subways in neighbouring areas) were positively associated with alcohol expenditures. A one-unit increase in subway intercepts is expected to increase expenditures by 5.01%. For lagged subways, a one-unit increase in lagged subway intercepts is expected to increase expenditures by 28.28%. Primary drinking restaurants and lagged primary drinking restaurants were both positively associated with alcohol expenditures. For primary drinking restaurants, a one-unit increase in primary drinking restaurant density by area is expected to increase expenditures by 0.06%. For lagged primary drinking restaurants, a one-unit increase in lagged primary drinking restaurant density by area is expected to increase expenditures by 0.73%. Finally, for restaurant density and lagged restaurant density, both variables were positively associated with alcohol expenditure. For restaurant density, a one-unit increase in restaurant density by area is expected to increase expenditures by 0.03%. For lagged restaurant density, a one-unit increase in lagged restaurant density by area is expected to increase expenditures by 0.10%.

## 4.3 Purchased in Stores Expenditure Data

### 4.3.1 Global Spatial Autocorrelation

To answer research question four, Global Moran's I was also applied to the purchased in stores expenditure to determine if geographic variation was also present in this dataset. For purchased in stores alcohol expenditure, the Global Moran's I value was 0.623331 and was significant at the $p = 0.001$ significance level (999 Permutations). As positive spatial autocorrelation values range from > 0 to +1, this value represented a significant and high degree of positive spatial autocorrelation. This result confirmed that there was significant geographic variation in the alcohol expenditure dataset by DA in the city of Toronto. Figure 4.3.1.1 presents the Moran's I scatterplot used to calculate the Global Moran's I value for the purchased in stores alcohol expenditure data. The purchased in stores expenditure values (x-axis) were plotted against expenditure values in nearby areas (y-axis) and the resulting positive slope indicated the presence of geographical variation.


Figure 4.3.1.1: Moran's I Scatterplot for Purchased in Stores Alcohol Expenditure: The Moran's I statistic plots the purchased in stores expenditure values (x-axis) against expenditure values in nearby areas (y-axis). The resulting correlation is the Moran's I value. Monte Carlo testing using 999 permutations was used to produce a p-value that determined whether the points plotted form a significant cluster. The Moran's I values were significant at the 0.001 significance level.

### 4.3.2 Local Spatial Autocorrelation – Hot-Spots and Cold-Spots

Having determined that there was geographic variation in the alcohol expenditure data, the next step was to answer the second half of research question four, which asked: Where were the locations of the hot-spots and cold-spots of alcohol expenditure? Figure 4.3.2.1 displays a significance map showing that six different cluster have been identified that are significant at the $p = 0.05$ significance level. Figure 4.3.2.2 identifies the clusters by hot-spot (red) and cold-spot (blue). There were a total of four hot-spots and two cold-spots identified for purchased in stores alcohol expenditures by using the Local Moran's I statistic.

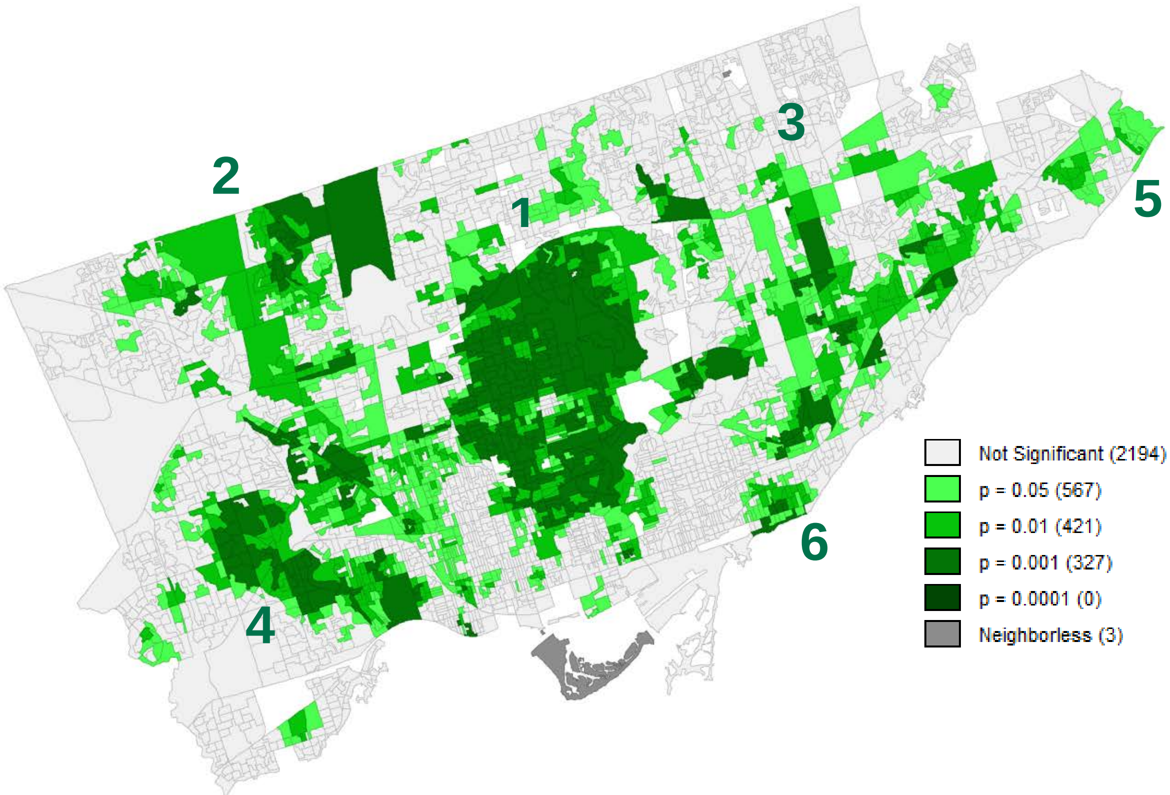



Figure 4.3.2.1: Significance Map for Clusters Identified for Purchased in Stores Alcohol Expenditures: This map shows the significance of the six clusters identified for purchased in stores alcohol expenditures. The darker the green colour, the more significant the DA is as part of the cluster.

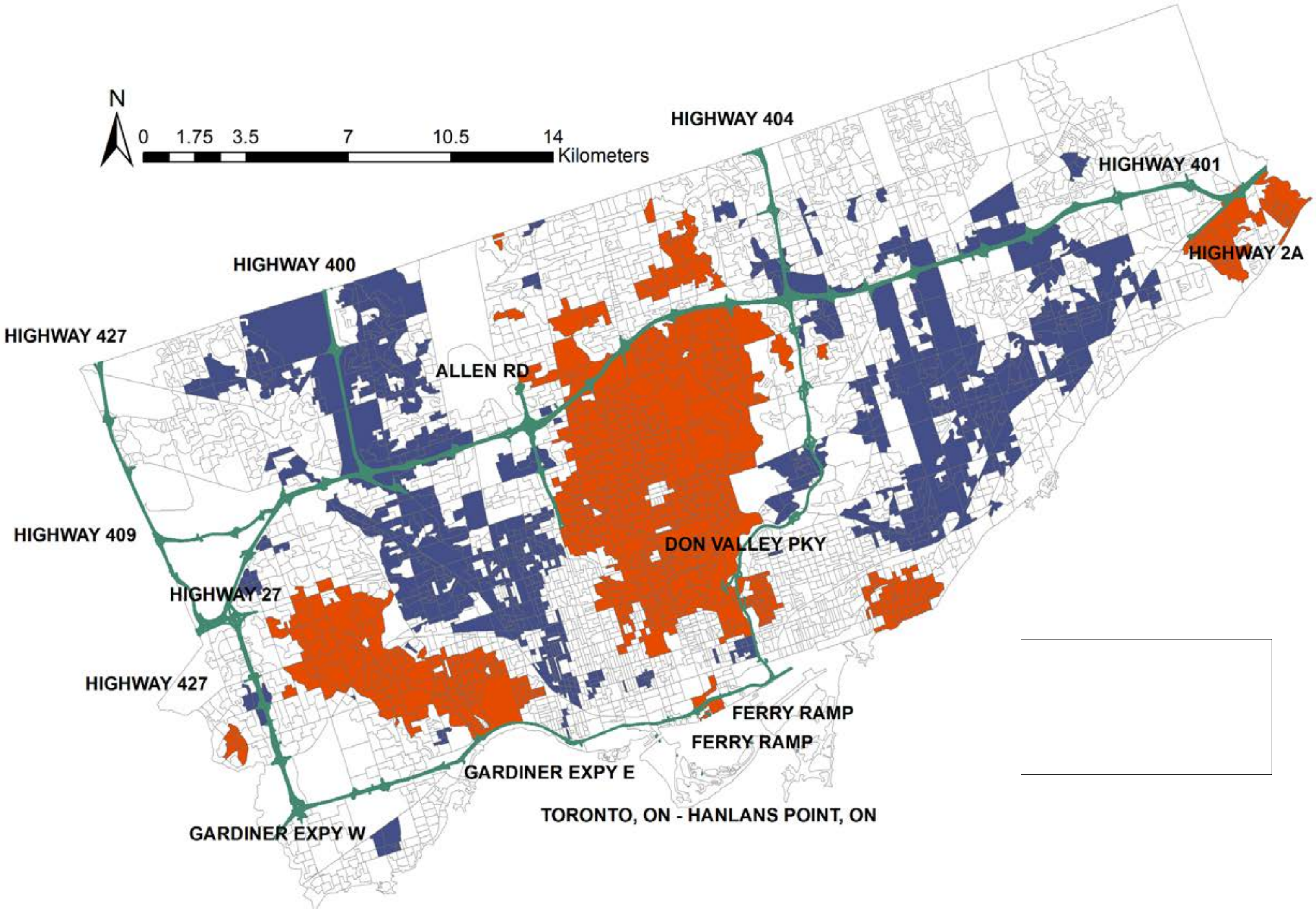
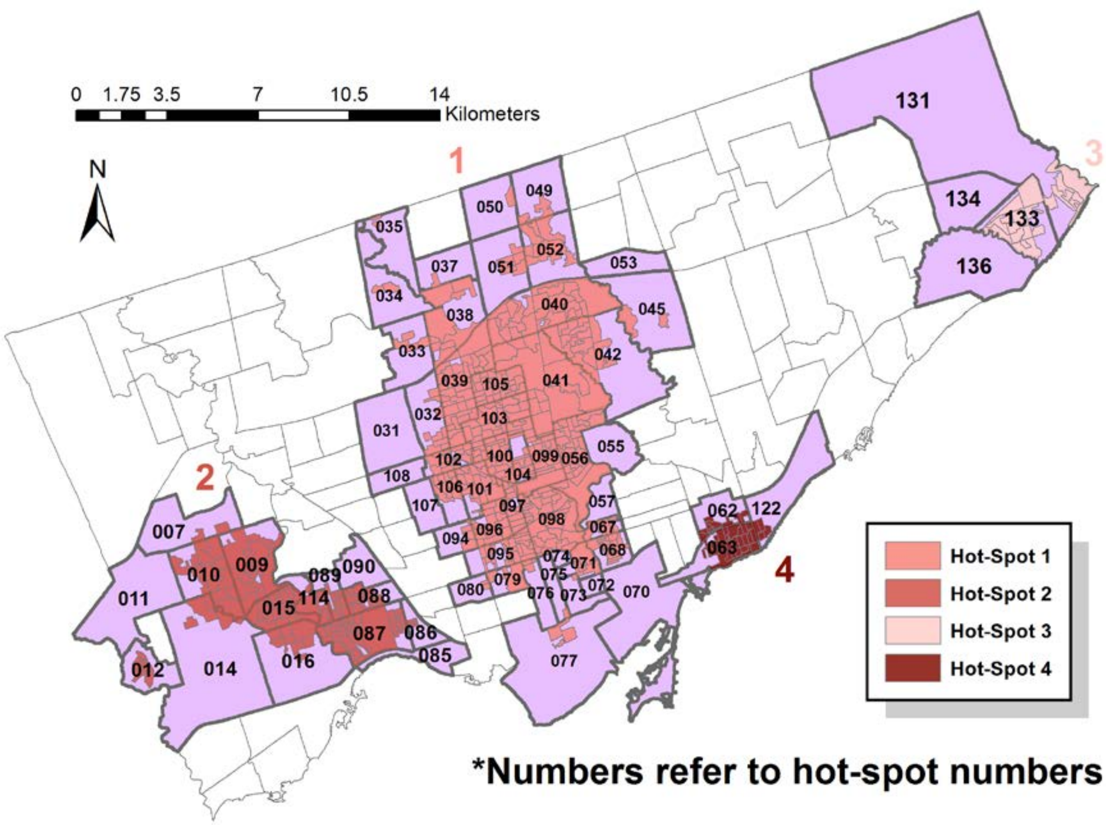



Figure 4.3.2.2: Hot-Spot and Cold-Spot Map for Purchased in Stores Alcohol Expenditures: This map shows the hot-spots (areas with high levels of alcohol expenditure clustered together) in red and the cold-spots (areas with low levels of alcohol expenditure clustered together) in blue. The major highways have also been overlaid to provide a sense of location for the clusters.

### 4.3.3 Changing Scales – Dissemination Area to Municipal Neighbourhoods

Once again, to extend the results of the analysis to policy-relevant neighbourhood area units, spatial selection methods were used to link the Dissemination Areas in hot-spots and cold-spots to corresponding municipal neighbourhoods. Figure 4.3.3.1 shows the municipal neighbourhoods corresponding to the hot-spot Dissemination Areas and Table 4.3.3-1 provides the summary information for the hot-spot clusters. Figure 4.3.3.2 shows the municipal neighbourhood corresponding to the cold-spot Dissemination Areas and Table 4.3.3-2 provides the corresponding summary information for the cold-spot clusters.




<u>Figure 4.3.3.1 – Hot-Spots with Corresponding Neighbourhoods: The DAs associated with hot-spots (in shades of red) were specifically extracted and overlaid with a map of the 140 municipal neighbourhoods. The neighbourhoods that are associated with these Dissemination Areas are highlighted in pink and are labeled with their neighbourhood number (in black).</u>

Table 4.3.3-1 – Hot-Spots for Expenditures from Purchases from Stores

| Cluster # | Type | North Boundary | East Boundary | South Boundary | West Boundary | Number of DA | Number of Neighbourhoods |
|---|---|---|---|---|---|---|---|
| **1** | Hot | Steeles Ave. | Victoria Park Ave. | Lake Ontario | Allen Rd./Dufferin Street | 433 | 47 |
| **2** | Hot | Dixon Rd/Hwy 401 | Keele St/Lansdowne Ave. | The Queensway | Etobicoke Creek | 116 | 15 |
| **3** | Hot | Hwy 401 | Scarborough-Pickering Townline | Lake Ontario | Morningside Ave. | 20 | 4 |
| **4** | Hot | Gerrard St E. | Warden Ave. | Lake Ontario | Coxwell Ave. | 37 | 3 |

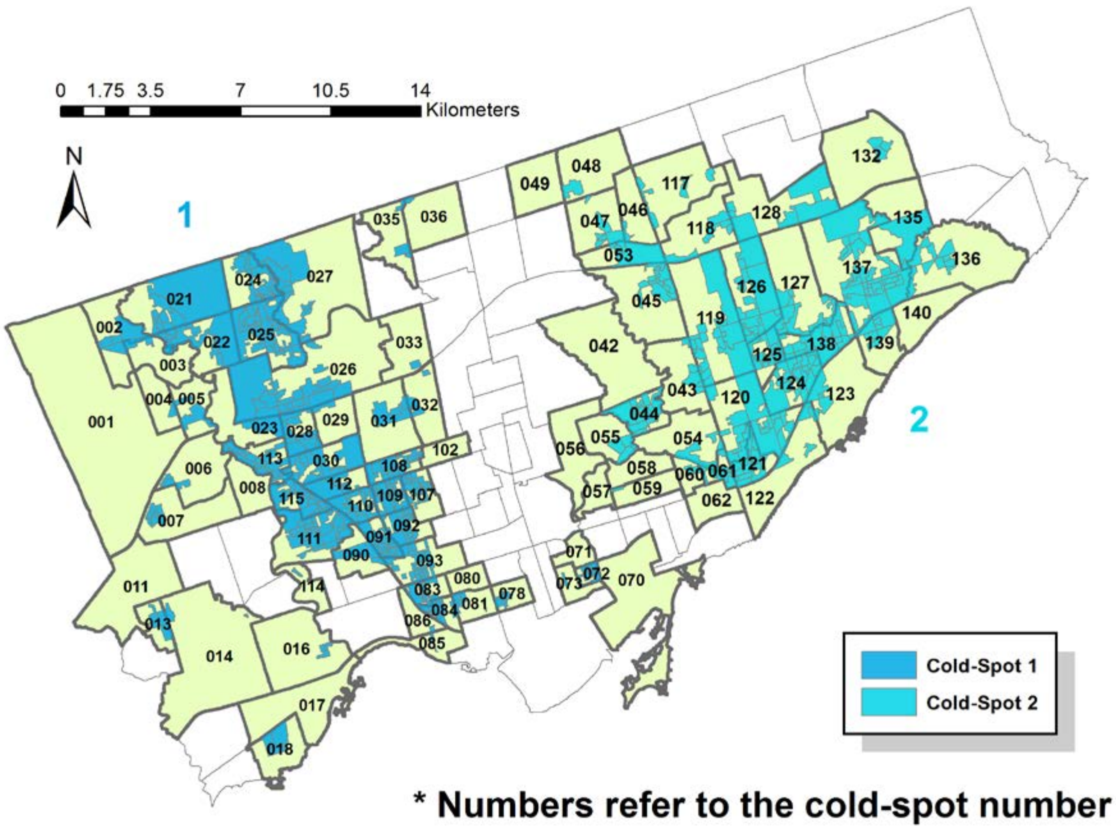



Figure 4.3.3.2 – Cold-Spots with Associated Neighbourhoods: The DAs associated with cold-spots (in teal and blue colours) were specifically extracted and overlaid with a map of the 140 municipal neighbourhoods. The neighbourhoods that are associated with these Dissemination Areas are highlighted in green and are labeled with their neighbourhood number (in black).

Table 4.3.3-2 - Expenditures from Purchases From Stores Cold-Spots - Summary

| Cluster # | Type | North Boundary | East Boundary | South Boundary | West Boundary | Number of DA | Number of Neighbourhoods |
|---|---|---|---|---|---|---|---|
| **1** | Cold | Steeles Ave | Avenue Rd./University Ave. | Lakeshore Blvd. West | Renforth Drive. | 382 | 54 |
| **2** | Cold | Steeles Ave | Meadowvale road | Queen st. east | Jarvis St. | 252 | 37 |

### 4.3.4 Spatial Regressions

The next step after determining the hot-spot and cold-spot locations was to examine the association between alcohol expenditure and the explanatory variables at the DA level in order to answer research question five. Research question five asked: What were the small-area level socio-economic and built-environment variables associated with alcohol expenditure? As the Global Moran's I measure indicated a significant amount of spatial autocorrelation, multivariate OLS regression models were no longer appropriate for the analysis. Four different spatial regressions models were used on the expenditure data and were then compared (Table 4.3.4-1). The best fitting model to answer research question five for purchased in stores expenditure was the spatial Durbin error regression model as it had the largest log-likelihood value, indicating the best fit (Table 4.3.4-1). The regression coefficients and significance of coefficients are provided in the table. The bivariate spatial error regressions can be found in Appendix A3.

Table 4.3.4-1: Spatial Regression Models for Purchased in Stores Alcohol Expenditures

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | **Purchased in Stores Expenditure** | | | |
| | ***spatial autoregressive*** | ***spatial error*** | ***Spatial Durbin*** | ***Spatial Durbin error*** |
| | **(1)** | **(2)** | **(3)** | **(4)** |
| **Socio-Economic Variables** | | | | |
| Percentage of 20 to 44 Year Olds | -0.057 | -0.390*** | -0.099 | -0.406*** |
| Percentage White | 0.003*** | 0.004*** | 0.003*** | 0.004*** |
| Percentage of Persons in White-collar Occupations | 0.364*** | 0.431*** | 0.353*** | 0.426*** |
| Median Household Income - After-Tax | 0.00001*** | 0.00001*** | 0.00001*** | 0.00001*** |
| Percentage of Income from Transfer Payments | -0.019*** | -0.018*** | -0.020*** | -0.018*** |
| Percentage of Apartments - 5 floors or more | -0.003*** | -0.003*** | -0.003*** | -0.003*** |
| Percentage of Apartments - Less than 5 floors | -0.002*** | -0.002*** | -0.002*** | -0.002*** |
| Percentage of Semi-Detached Houses | -0.002*** | -0.002*** | -0.002*** | -0.002*** |
| Percentage of Single Detached Houses | -0.002*** | -0.001** | -0.002*** | -0.001** |
| Percentage of Rented Dwellings | -0.001** | -0.001*** | -0.001** | -0.001*** |
| Percentage of Lone-Parent Families | -0.004*** | -0.002*** | -0.004*** | -0.002*** |
| Percentage Post-Secondary Education | 0.257*** | 0.230*** | 0.239*** | 0.219*** |
| **Built-Environment Variables** | | | | |
| Outlet Density by Area (KM - Squared) | 0.002 | 0.001 | 0.001 | 0.002 |
| Lagged Outlet Density by Area (KM - Squared) | - | - | 0.008** | 0.006 |
| Subway Intercepts | 0.056** | 0.052* | 0.017 | 0.059** |
| Lagged Subway Intercepts | - | - | 0.101** | 0.200*** |
| Constant | 4.005*** | 5.874*** | 4.035*** | 5.871*** |
| **Spatial Autoregressive Term** | | | | |
| Rho | 0.299*** | - | 0.298*** | - |
| Lambda | - | 0.654*** | - | 0.645*** |
| **Log Likelihood** | -542.482 | -424.912 | -533.770 | -414.403 |
| *Note:* | | | | *p<0.05; **p<0.01; ***p<0.001 |

### 4.3.5 Final Diagnostics – Testing for Remaining Spatial Dependence

Research question six asked: Is there any remaining geographical variation in the expenditure after controlling for socio-economic and built-environment factors? The final step of the analysis, as well as to answer research question six, was to determine whether there was any remaining spatial autocorrelation in the model or if the spatial regression model had accounted for it all. The residuals for the purchased in stores spatial Durbin error model were tested with Moran's I and produced a value of -0.047. Figure 4.3.5.1 presents the Moran`s I scatterplot of the residuals. Similar to licensed premises expenditures, the Moran's I value for the spatial Durbin error model once again represented a significant decrease from the Moran's I value of the linear model residuals (0.272). As the value was very close to zero, this also indicated that the spatial autoregressive error term had eliminated the spatial autocorrelation originally present in the data. As a result, this meant that geographical variation in the data no longer presented an issue and proved that the spatial regression model had performed as expected.

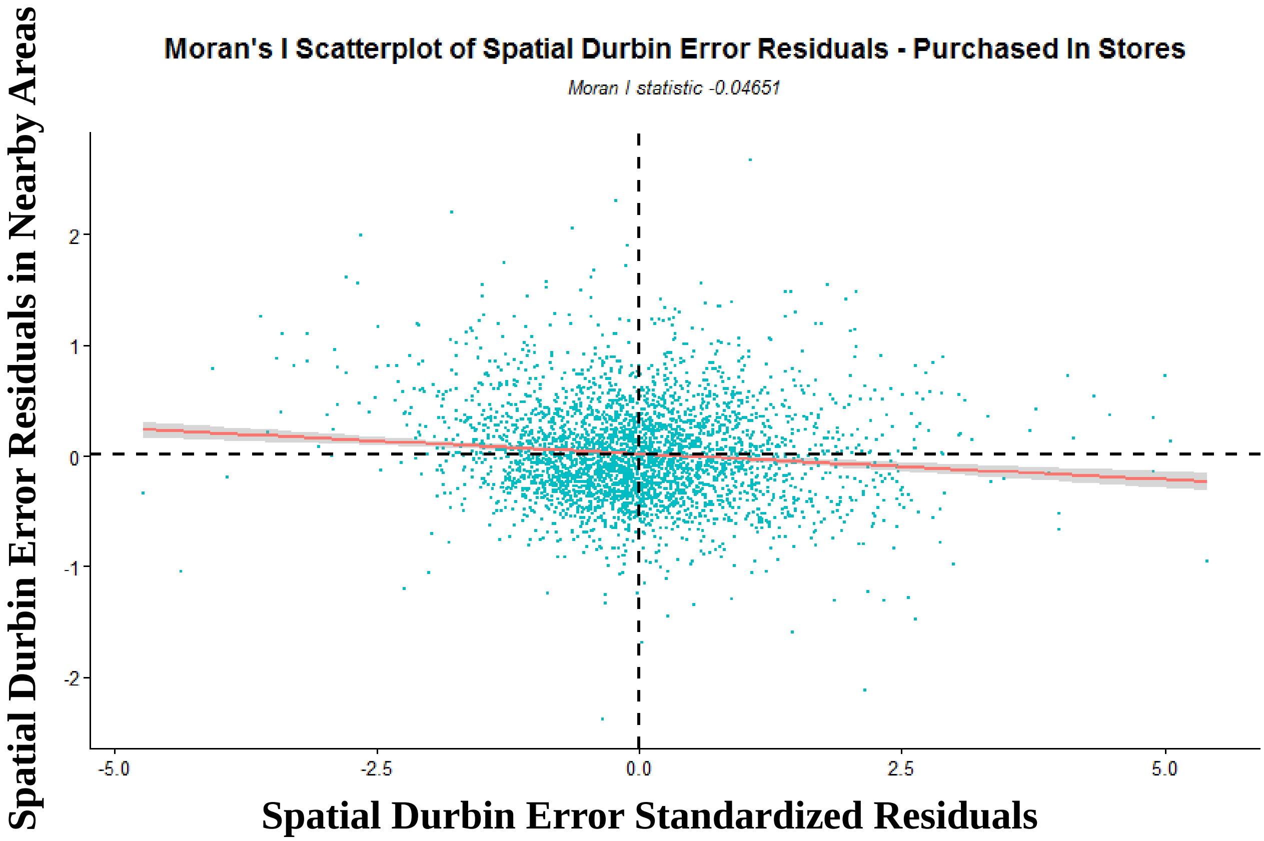



Figure 4.3.5.1 - Moran's I Scatterplot of SDEM Residuals for Purchased In Stores Expenditures: The standardized residuals for the SDEM model (x-axis) are plotted against the spatially lagged standardized residuals for the SDEM model. The resulting slope is the Moran's I value. The value of -0.047 indicates that geographical variation is no longer affecting the model as it is very close to zero.

### 4.3.6 Interpreting the Coefficients for the Spatial Durbin Error Model – Purchased in Stores Alcohol Expenditures

#### *4.3.6.1 Socio-economic Variables*

For age measures, the percentage of 20 to 44 year old was negatively associated with alcohol expenditures. A one-unit increase in percentage of 20 to 44 year olds is expected to decrease expenditures by 33.36%. For ethnicity measures, percentage White was the only significant variable and was positively associated with alcohol expenditures. A one-unit increase in percentage White is expected to increase alcohol expenditures by 0.40%. Occupation as measure by percentage of white-collar occupations was positively associated with alcohol expenditures. A one-unit increase in percentage of white-collar workers is expected to increase expenditures by 53.07%. Median after-tax incomes was significantly positively associated with alcohol expenditures, while conversely, social assistance measured as percentage of income from transfer payments was negatively associated with alcohol expenditures. For median after-tax income, a one-unit increase in median after-tax income is expected to increase expenditures by 0.0006%. For transfer payments, a one-unit increase in percentage of income from transfer payments is expected to decrease expenditures by 1.81%. For housing measures, all five significant housing variables were negatively associated with alcohol expenditures. For apartments with five floors or more, a one-unit increase in the percentage of apartments with five floors or more is expected to decrease expenditures by 0.30%. For apartments with five floors or less, a one-unit increase in the percentage of apartments with less than five floors is expected to decrease expenditures by 0.16%. For semi-detached houses, a one-unit increase in the percentage of semi-detached houses is expected to decrease expenditures by 0.17%. For single-detached houses, a one-unit increase in the percentage of single-detached houses is expected to decrease expenditures by 0.11%. Lastly, for percentage of rented dwellings, a one-unit increase is expected to decrease expenditures by 0.13%. Lone-parents households were also found to be negatively associated with alcohol expenditures. A one-unit increase in lone-parent headed households is expected to decrease expenditures by 0.23%. Education as measured by percentage with post-secondary degree was positively associated with alcohol expenditures. A one-unit increase in percentage of post-secondary education is expected to increase expenditures by 24.44%.

#### *4.3.6.2 Built-Environment Variables*

Subways and lagged subways were the only significant built-environments associated with alcohol expenditures. Both of these variables were positively associated with alcohol expenditures. For subways, a one-unit increase in subway intercepts is expected to increase expenditures by 6.03%. For lagged subways, a one-unit increase in lagged subway intercepts is expected to increase expenditures by 22.14%.

## 4.4 Comparison of Best Models

To examine the differences between the licensed premises and purchased in stores results, Table 4.3.6-1 provides a comparison table featuring the best model for the licensed premises expenditure data and the best model for the purchased in stores expenditure data. The final SDEM regression model for licensed premises expenditure had nine socio-economic variables, seven built-environment variables and the lambda variable for a total of seventeen explanatory variables. The final SDEM regression model for purchased in stores expenditure data had twelve socio-economic variables and five built-environment variables and the lambda variable for a total of eighteen explanatory variables.

Table 4.3.6-1 - Comparison Table for Best Models

| Spatial Regressions for both Expenditures | | |
|---|---|---|
| | *Dependent variable:* | |
| | **Licensed Premises Expenditure** | **Purchased From Stores Expenditure** |
| | **(1)** | **(2)** |
| **Socio-Economic Variables** | | |
| Percentage of 20 to 44 Year Olds | | -0.406*** |
| Percentage White | 0.004*** | 0.004*** |
| Percentage of Filipino Ethnicity | -0.003* | |
| Percentage of Persons in White-collar Occupations | 0.580*** | 0.426*** |
| Median Household Income - After-Tax | 0.00001*** | 0.00001*** |
| Percentage of Income from Transfer Payments | -0.018*** | -0.018*** |
| Percentage of Apartments - 5 floors or more | | -0.003*** |
| Percentage of Apartments - Less than 5 floors | 0.001*** | -0.002*** |
| Percentage of Semi-Detached Houses | | -0.002*** |
| Percentage of Single Detached Houses | 0.002*** | -0.001** |
| Percentage of Row Houses | 0.005*** | |
| Percentage of Rented Dwellings | | -0.001*** |
| Percentage of Lone-Parent Families | | -0.002*** |
| Percentage Post-Secondary Education | 0.300*** | 0.219*** |
| **Built-Environment Variables** | | |
| Outlet Density by Area (KM - Squared) | | 0.002 |
| Lagged Outlet Density by Area (KM - Squared) | - | 0.006 |
| Subway Intercepts | 0.049 | 0.059** |
| Lagged Subway Intercepts | 0.249*** | 0.200*** |
| Primary Drinking Restaurant Density by Area (KM - Squared) | 0.001 | - |
| Lagged Primary Drinking Restaurant Density by Area (KM - Squared) | 0.007* | - |
| Restaurant Density (KM - Squared) | 0.0003* | - |
| Lagged Restaurant Density (KM - Squared) | 0.001** | - |
| Constant | 4.506*** | 5.871*** |
| **Spatial Autoregressive Error Term** | | |
| Lambda | 0.639*** | 0.645*** |
| | | |
| *Note:* | | *p<0.05; **p<0.01; ***p<0.001 |

For both expenditure types, the SDEM model provided a significant improvement in log-likelihood value compared with the other spatial models. The $\lambda$ error term also provides important information and diagnostic regarding the model. As the $\lambda$ error term is highly significant in both models, this indicates that there is significant spatial clustering as well as the possibility of other important explanatory variables being missing in the model.

# Chapter 5: Discussion

The main objectives of this study were to identify and quantify the geographical variation in two types of alcohol expenditure data in the City of Toronto, identify the hot-spots and cold-spots of alcohol expenditure in Toronto, and identify small-area socio-economic and built-environment determinants for alcohol expenditure at the Dissemination Area Level. Following the discussion of the findings, this chapter discusses the limitations and strengths that stem from applying spatial methods to study alcohol expenditure data for the first time. This chapter concludes with a discussion of the public health implications as a result of the findings from this study and future research directions.

## 5.1 Global Spatial Autocorrelation and Local Spatial Autocorrelation

### 5.1.1 Licensed Premises Geographic Variation

The licensed premises expenditure data was expected to show geographical variation as it was expected that the expenditure patterns would vary according to socio-economic factors, which are spatially distributed across the City of Toronto. To identify geographical variation, the expenditure data was first mapped to a choropleth map as shown in Figure 5.1.1.1.

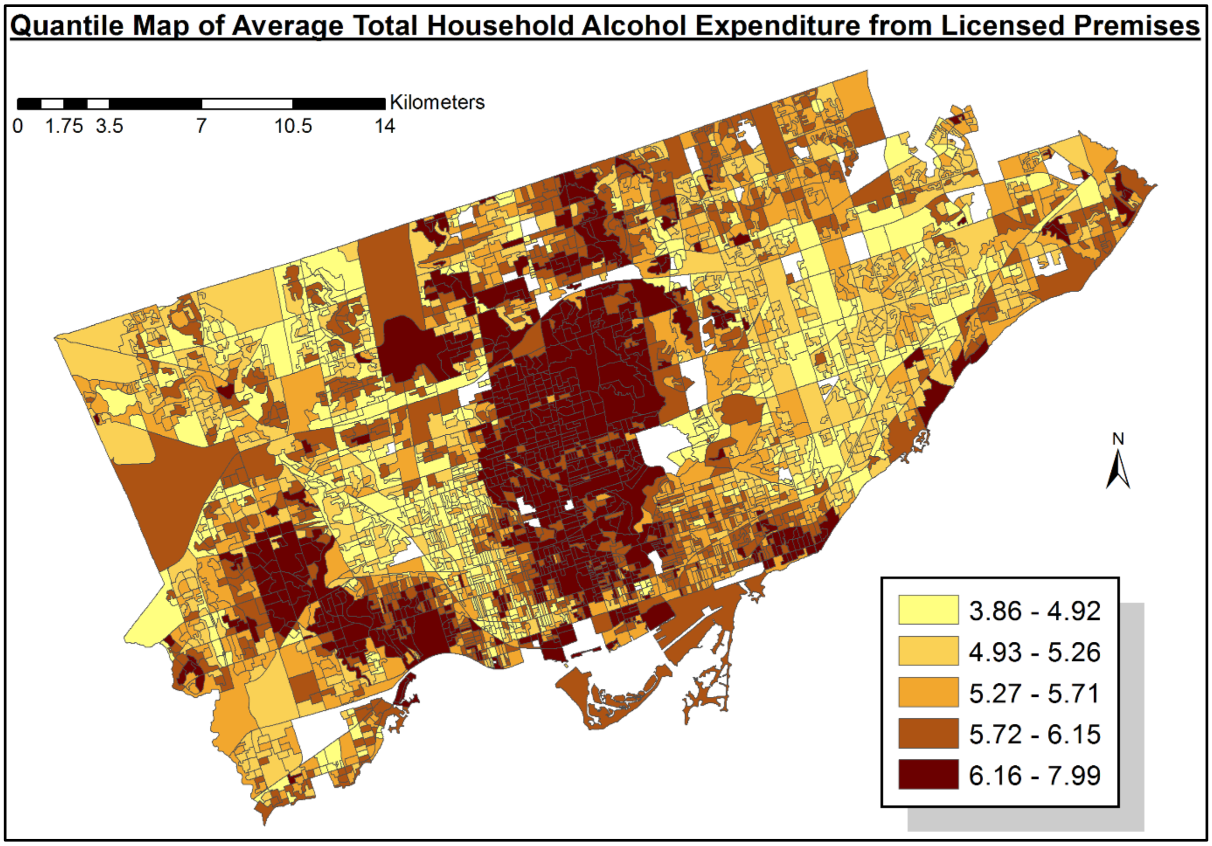



Figure 5.1.1.1 Quantile Map of Licensed Premises Aclohol Expenditures: This map displays Dissemination Areas by their licensed premises expenditure data using a five quantile classes. Darker colours indicate a higher quantile; therefore, yellow represents the lowest level of alcohol expenditures (lowest quantile) and dark brown represents the highest level of alcohol expenditures (highest quantile).

From a visual examination of the quantile map of licensed premises expenditure in Figure 5.1.1.1, it appeared that similar values of alcohol expenditure from licensed premises were clustered together. In order to confirm this, Global Moran's I value was calculated to quantify the average level of spatial autocorrelation in the overall dataset. The resulting Global Moran's I value for licensed premises expenditures indicated that there was significant geographical variation in the expenditure data. As such, there was a high likelihood of identifying distinct clusters of alcohol expenditure through the use of local clustering methods.

Application of local clustering methods to expenditure data revealed three hot-spots and three cold-spots specifically, as seen in Figure 4.2.2.2. For the hot-spots, the first cluster was the largest hot-spot and spanned an area from Steeles avenue in North York, through the centre of the city to the lakeshore in Downtown Toronto. The second cluster was spread diagonally across the south-west of Toronto to the west of Toronto. Using the historical boroughs for geographical context, this coincided with downtown Toronto and Etobicoke. This was also the second largest cluster. Finally, the third and smallest cluster was located in the south-east of Toronto, bordering Lake Ontario and the south-eastern edge of downtown Toronto.

In terms of the cold-spots, the first cold-spot was the largest of the three clusters and was located in the east of Toronto. This cluster included a large portion of Scarborough, along with small portions of both North York and the downtown Toronto area. The second cold-spot was located in the south-west end of Toronto, starting south of the 401, and extending in a south easterly fashion towards the south of Toronto. This cold-spot encompassed Etobicoke, downtown Toronto, as well as a small portion of North York. The third and smallest cold-spot was located in the north-west of Toronto and was spread across Etobicoke, with peripheral parts in North York.

Figure 4.2.2.1 provides a map demonstrating the significance of each cluster. The map indicates that the cores of each of the clusters are significant at $p<0.001$, while the peripheral regions of each cluster are slightly less significant at $p<0.01$.

### 5.1.2 Purchased in Stores Geographic Variation

Similar to licensed premises expenditures, alcohol expenditures from purchases in store were hypothesized to exhibit geographical variation. The choropleth map in Figure 5.1.2.1 was produced to help identify geographic variation in the purchased in stores alcohol expenditures.

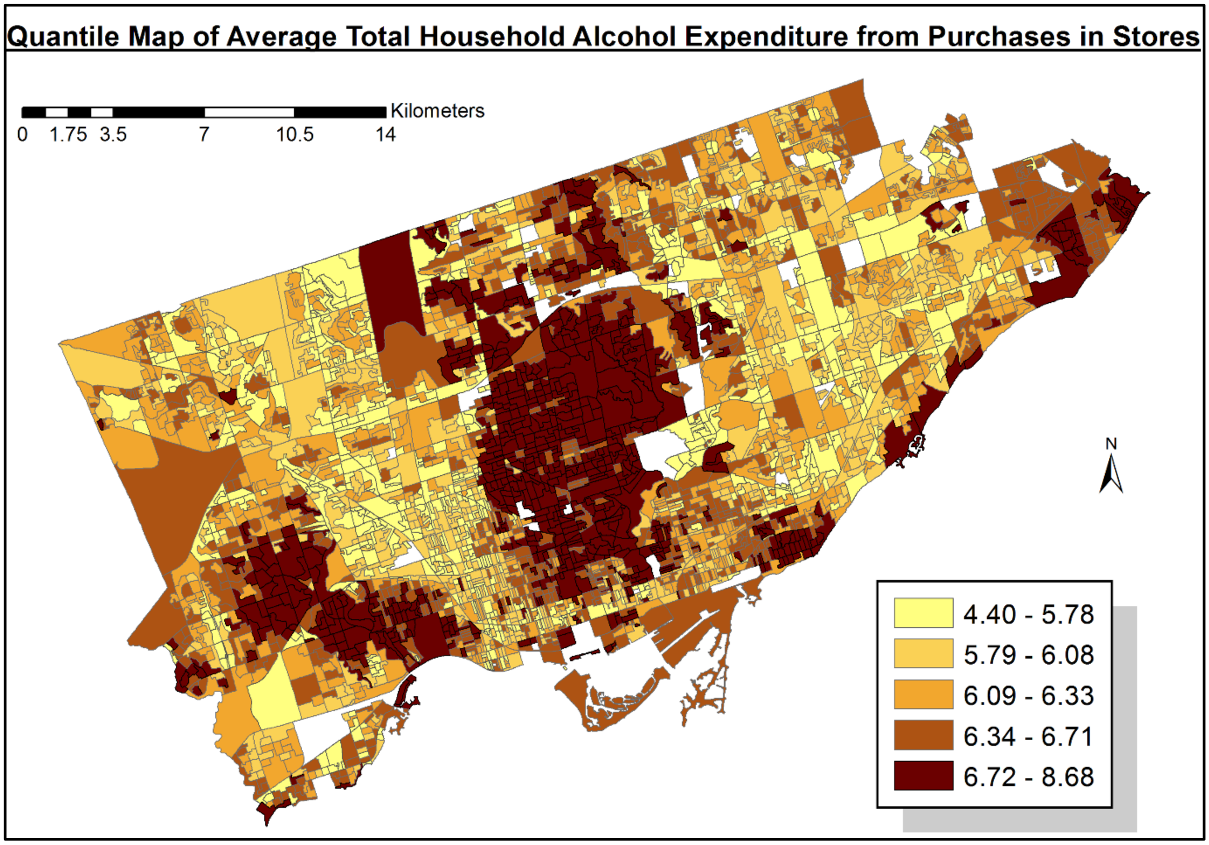



Figure 5.1.2.1: Quantile Map of Purchased in Stores Alcohol Expenditures: This map displays Dissemination Areas by their purchased in stores expenditure data using a five quantile classes. Darker colours indicate a higher quantile; therefore, yellow represents the lowest level of alcohol expenditures (lowest quantile) and dark brown represents the highest level of alcohol expenditures (highest quantile).

Visual examination of the quantile map of alcohol expenditures from purchases in stores in Figure 5.1.2.1 again suggested that areas with similar levels of alcohol expenditure were clustered together. This was confirmed by the Global Moran's measure, which indicated that there was significant geographical variation in the expenditure data. As such, there was once again a high likelihood of identifying distinct clusters of alcohol expenditure through the use of local clustering methods.

As seen in Figure 4.3.2.2, six different clusters were again identified for expenditures from purchases in stores; however, there were four hot-spots and only two cold-spots, compared to three of each in the licensed expenditure data. The first and largest hot-spot extended across the centre of the city - running south of Highway 401 down towards the lakeshore. The hot-spot also included portions of North York as well as downtown Toronto. The second hot-spot was located in the south-west of Toronto, covering a diagonal area from Highway 427 to the Gardiner Expressway East. This hot-spot included large portions of Etobicoke and portions of downtown Toronto. The third hot-spot was located in the very east of the city, with the entire cluster residing within Scarborough and bordering the Scarborough-Pickering townline. The fourth and smallest hot-spot was located in the south-east of Toronto on the eastern border of downtown Toronto.

The first of the two cold-spots was located in the west of Toronto and extended southwards from Steeles towards the Gardiner Expressway. This was the larger of the two cold-spots and included a sizeable portion of Etobicoke, along with smaller parts of North York and downtown Toronto. The second cold-spot was located in the east of the city, with the majority lying to the east of Highway 404. The cold-spot was comprised of a large portion of Scarborough, along with smaller parts of North York and downtown Toronto.

Figure 4.3.2.1 illustrates the significance level of each cluster. With the exception of the hot-spot in the east end, the map indicated that the cores of each of the clusters were significant at $P<0.001$, while the peripheral regions of each cluster were slightly less significant at $p<0.01$.

The hot-spot at the very east end had a core that was significant at the $p<0.01$ significance level, with peripheral regions significant at the $p<0.05$ significance level.

For the final part of the cluster analysis, the neighbourhood map of Toronto was overlaid onto the hot and cold-spots for each expenditure type, and the neighbourhoods associated with each cluster were then delineated using a spatial selection function in ArcMap. This provides additional geographical context by helping to link the Dissemination Area level expenditure findings to the neighbourhood level in Toronto. Using the Dissemination Areas allows highly localized areas of high and low expenditure to be detected at the small-area level; however, as Dissemination Areas are standardized for population, they do not necessarily reflect real neighbourhood contexts. The one hundred and forty neighbourhoods were created to evaluate socio-economic variables within the city in a meaningful geographical context. As such, they are widely used by community agencies for planning purposes, service provision and longitudinal data acquisition. Therefore, these neighbourhood units may provide area units that have greater social significance and practical implications. As the alcohol expenditure data has been mapped from the Dissemination Area level to the neighbourhood level, future studies can examine alcohol expenditure using this neighbourhood scale. Future studies can also take advantage of the longitudinal data and in-depth socio-economic profiles available for these neighbourhoods to look at how these factors correspond to changes in alcohol expenditure over time.

## 5.2 Socio-Economic and Built-Environment Characteristics Associated with Licensed Premises and Purchased in Stores Alcohol Expenditures

### 5.2.1 Socio-Economic Characteristics Associated with Licensed Premises and Purchased in Stores Alcohol Expenditures

The hypotheses linking each explanatory variable to alcohol expenditure were kept the same for both expenditure types, as previous expenditure literature does not differentiate between the locations of purchase. The following section details whether the socio-economic variables examined for both types of expenditures were significant or insignificant. The hypothesis for each explanatory variable is provided, along with a discussion of whether the results were in agreement with it.

The percentage of 18 to 44 year olds within each DA was expected to have a significant positive effect on alcohol expenditures as the prime consumption age for males is 25 to 49 and 18 to 24 for females (19, 33, 56). For Canada, heavy drinking is also most prevalent between the ages of 18 to 34 (51). For the licensed premises expenditures however, the percentage of 18 to 44 year olds within each DA was not significant. This interesting finding mirrored the one in the Malaysian study by Tan, Yen (32) which did not find significant differences in alcohol expenditures due to age. There are several possible reasons for this finding. The first possibility is that that at the Dissemination Area level, population counts are too low to allow the effects of specific age groups to reach significance. It is also possible that despite consuming greater volumes of alcohol, the 18 - 44 year old age group may is still being outspent by older age groups, who may be able to purchase alcohol that is more expensive. In support of this, Tan, Yen and Nayga (32) found that households with younger age groups (15 to 30) spent considerably less than their middle aged counterparts (31 to 56 years) after stratifying household expenditures by ethnicity. If stratified expenditure data is available in the future, this avenue should be explored. Surprisingly, the percentage of 18 - 44 year olds in each DA was determined to have a significant negative association with purchased in store expenditure. Once again, a potential explanation for this is that while the 18 to 44 year old age group consumes the greatest volume of alcohol, it may be mostly inexpensive alcohol, resulting in lower total expenditures compared to other age groups. Accordingly, future research should examine the middle aged and senior age groups for associations with alcohol expenditures.

With regards to gender, a greater ratio of males to females was expected to have a positive effect on level of expenditure. This hypothesis was in keeping with previous studies, which showed that when compared to females, males were more likely to spend more on alcohol and consume greater quantities (15, 19, 30, 32). Here, however, a greater ratio of males to females at the DA level was not significantly association with alcohol expenditure. One possible reason for this discrepancy is that the counts of males and females at the DA level may not be sufficient to provide a sizeable male-to-female ratio that would demonstrate stronger associations with alcohol expenditure. Another possibility is that over time, the difference in alcohol expenditure between males and female may be decreasing. This would be consistent with alcohol use trends, which show that male and female levels of alcohol use are converging (30). Future studies should examine the effect of the ratio of males to females at other spatial scales on alcohol expenditures.

In terms of ethnicity, it was expected that the percentage of the population identifying as White would be positively associated with alcohol expenditures. Several other major minority groups and their respective percentages within the population were also expected to be associated with alcohol expenditures as well. As hypothesized, the percentage of the population identifying as White was significantly associated with increased licensed premises and purchased at store alcohol expenditures. This was in keeping with previous work by Yen and Jenson (33) which found that a White head of household was associated with increased alcohol expenditures. However, among the minority groups, only the percentage of the population identifying as Filipino demonstrated a significant association with alcohol expenditures. Specifically, the percentage of Filipino people had a significant negative association with licensed premises alcohol expenditure. Previous studies have shown differences in levels of alcohol expenditures between various ethnic groups in different countries as well as within the same country (19, 32). As such, it was expected that percentages of major minority groups - including Chinese, Black and South Asian - would be significantly associated with alcohol expenditure. A possible explanation for the lack of associations is that the pooling of data may have masked differences between ethnicities. Tan, Nayga and Yen (32) first used pooled expenditures to look at socio-economic associations with alcohol expenditures before using alcohol expenditure data separated by ethnicities to find differences between the major ethnic groups in Malaysia. If available, future studies should follow this template and use alcohol expenditure data stratified by ethnicity to look for significant differences in alcohol expenditure levels between the major ethnicities in Toronto.

White-collar occupations were hypothesized to be positively associated with both types of alcohol expenditures. This was supported by the work of Yen and Jensen (33) as well as that of Tan, Nagaya and Yen (32), who found that that white-collar occupations were positively associated with alcohol expenditures in the US and Malaysia respectively. While white-collar occupations were indeed found to be significantly positively associated with alcohol expenditures, it was also seen that higher median household income had only a slight positive association with alcohol expenditures. It is interesting that such discrepancy in the strength of associations should exist, given that white-collar occupations are usually higher paying. This may indicate that the impact of occupation on alcohol expenditure has more to do with the drinking culture of the occupation as opposed to strictly the income. This reasoning is plausible as Sharpe et al. (19) found a significant drinking culture in certain industries in Korea. Additional studies should focus on looking at different types of white-collar occupations to determine if specific occupations would be positively or negatively associated with alcohol expenditures.

Median after-tax household income was hypothesized to be positively associated with both types of alcohol expenditures. This expectation was consistent with the findings of previous studies that showed household incomes had a significant positive effect on alcohol expenditures (15, 32, 33), As aforementioned mentioned, while median after-tax income was found to be positively associated with both alcohol expenditure types,  the overall effect was miniscule in comparison to what was expected. Therefore, while the effect of income was significant, it was not a major contributing factor to alcohol expenditure.

Social assistance was expected to be negatively associated with both alcohol expenditure types. As social assistance is an indicator of low-income families with less disposable income, it was expected to have the reverse effect of high median after-tax income. The results confirmed this hypothesis as social assistance in the form of percentage of income from government transfer payments was determined to have an inverse relationship with both types of alcohol expenditures. It should also be noted that while this was the expected finding, this finding contradicts the one by Abdel-Ghany and Silver (15), which had previously found that social assistance was not a significant predictor for alcohol expenditures in Canada. A possibility for this discrepancy is that the Abdel-Ghany and Silver considered household expenditures from the entire country, whereas this study focused specifically on the study region of Toronto. Local differences between rural areas and urban centres may have been masked at the spatial scale used in their study. Future studies should examine other Canadian urban centres and rural areas to see if the association between social assistance and alcohol expenditure varies according to the study region.

All of the housing variables were hypothesized to be positively associated with both alcohol expenditure types. It was expected that these housing measures would be indicative of areas with higher residential population and therefore, would be serviced by a greater number of licensed premises, consequently leading to higher expenditures. For licensed premises expenditure, three housing variables were significant: percentage of apartments with less than five floors, percentage of single-detached houses and row houses. All three were significantly positively associated with alcohol expenditures, which seems to support this hypothesis. For purchased in stores expenditures four housing variables were significant: percentage of apartments with five or more floors, percentage of apartments with less than five floors, percentage of semi-detached houses and percentage of single-detached houses. However, contrary to the licensed premises expenditure, all four housing variables were negatively associated with purchased in stores expenditures. This was unexpected and contrary to the expected hypothesis. A potential cause is that increased housing of any type is associated with a greater residential population and a reduced area for alcohol outlets due to the small area of Dissemination Areas. Future studies should explore reasons for differences in alcohol expenditure behaviour between families in different housing types.

It was hypothesised that percentage of rented dwellings would be negatively associated with both alcohol expenditure types, since increased percentage of rented dwellings correlates with decreased homeownership and consequently, greater housing insecurity, leading to less available income to spend on alcohol. The results however, showed that percentage of rented dwellings was not significantly associated with either type of alcohol expenditures. This finding mirrored the one by Yen and Jensen (33), which looked at homeownership as a proxy for assets and also found that the association was not significant.

Marital status or being married was hypothesized to have a negative association with both alcohol expenditure types. Previous studies had shown a discrepancy in findings, with Abdel Ghany and Silver (15) finding no significant association between being married and alcohol expenditures whereas Yen and Jenson (33) found that married couples were negatively associated with alcohol expenditures. Sharpe et al. (19) also found that married couples did not have a significant effect on alcohol expenditures and consumption decisions. Therefore, an inverse relation was thought to be most likely type of significant association. In the end however, marital status was not significantly associated with either alcohol expenditure type. These findings agreed with the Abdel Ghany and Silver and Sharpe et al. studies.

It was hypothesized that since lone-parent households would be negatively associated with alcohol expenditures, as lone-parent households are a measure of deprivation due to the greater financial burden of a single head of household supporting dependants. This hypothesis was partially confirmed as lone-parent households were determined to be an insignificant predictor of licensed premises expenditures but had a negative association with purchased in stores expenditures.

It was hypothesized that percentage with post-secondary education would be positively associated with both expenditure types. Previous studies have shown differences in association depending on the location of the study. In studies conducted in Western developed countries, high education is associated with increased alcohol expenditure whereas in Asia, high education has been found to decrease alcohol expenditures (19). Despite this discrepancy, it was hypothesized that the results from this study would be similar to those performed in other developed Western nations. The results were in agreement as percentage with post-secondary education was determined to have significant positive association with both types of alcohol expenditures. After occupation, percentage with post-secondary education had the largest impact on both types of alcohol expenditures. It was also hypothesized that increased educational attainment would result in higher occupational status and greater income, leading to higher alcohol expenditure levels. However, as median after-tax income was found to have only a nominal positive effect on alcohol expenditure, it is possible that these results are indicative that a drinking culture at post-secondary institutions might be the cause of the differences between Western and Eastern countries. Future research should investigate how post-secondary settings in the Western and Eastern counties lead to differences in alcohol expenditure behaviour.

Average number of persons in private households was hypothesized to be negatively associated with alcohol expenditures. An increase in the number of persons in a household was expected to result in less disposable income to spend on luxury goods such as alcohol (32). Previous studies had shown this negative association (32, 125). Here, however, average number of persons in private households was not significantly associated with either type of alcohol expenditure.

Aboriginal identity was hypothesized to be positively associated with alcohol expenditures due to the elevated prevalence of alcohol consumption among the aboriginal population (77). It was determined however, that aboriginal identity was not significantly associated with either type of alcohol expenditure. A possible reason for the lack of association is that the aboriginal population maybe too small to be accurately captured at the Dissemination Area level. Future research should examine aboriginal identity and alcohol expenditure at different spatial scales for possible associations.

### 5.2.2 Built-Environment Characteristics Associated with Licensed Premises and Purchased in Stores Alcohol Expenditures

Proportion of residential land and proportion of commercial land were both not significant for both types of alcohol expenditures. Both land use types were expected to be positively associated with alcohol expenditures as an increased number of residents due to increased residential land and increased number of alcohol outlets due to commercial land, was expected to contribute to increased alcohol expenditures. For residential land, there is the possibility that an increased number of residents does not equate to a greater availability of licensed premises or alcohol outlets in the same DA as the DAs are small areas. Therefore, an increased residential population does not equate to increased alcohol expenditures. For commercial areas, it appears that it may not be the best proxy for licensed premises or alcohol outlets, as the amount land that is designated as commercial areas in Toronto does not appear to correspond with where licensed premises exist. As seen in Figure 5.2.2.1, the restaurant density and primary drinking restaurant density variables were found to be much more accurate in terms of representing the actual amount of licensed premises in Toronto.

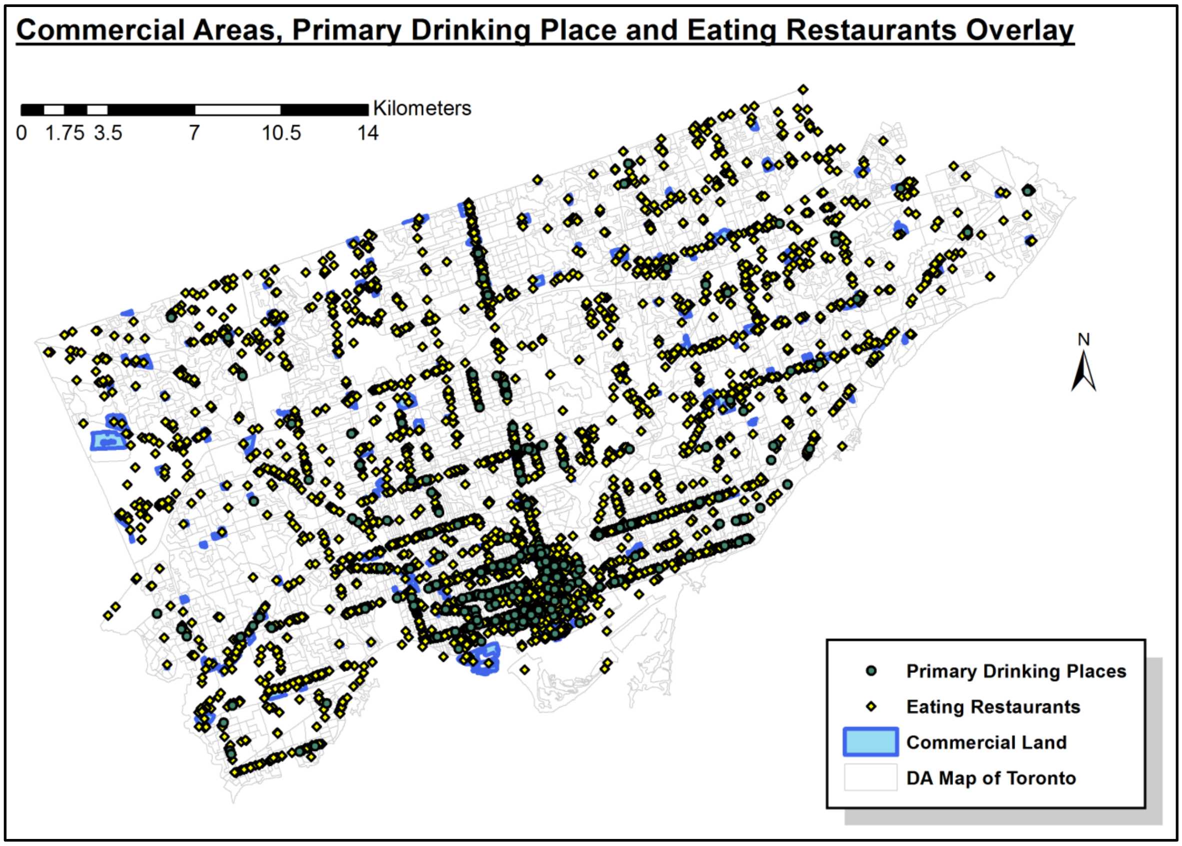



Figure 5.2.2.1: Map of all Primary Drinking Restaurants, Eating Restaurants and Commercial Land Overlaid on Toronto DA Map : Commercial land is represented by the areas in light blue with darker blue boarders. Primary drinking restaurants are represented by green circles and eating restaurants were represented by yellow diamonds. This map shows the discrepancy between the commercial areas and drinking and eating restaurants. This suggests that licensed premises are better represented by the primary drinking and eating restaurants as compared to commercial land use.

This was the first study to identify a positive association between alcohol expenditure and subway intercepts. Subway intercepts and lagged subway intercepts were considered to be measuring the same variable due to the small size of each Dissemination Area. It is unlikely that a Dissemination Area would have access to a subway line due to the small size of a DA; however, it is much more likely that a DA might boarder a nearby DA that had access to a subway line (measured by lagged subway intercepts). Therefore, subway access was examined using these two variables. Overlay analysis in Figure 5.2.2.2 showed that the areas of high alcohol expenditure that formed the large cluster running through downtown Toronto to North

York was surrounded by subway lines. It was hypothesized that access to the downtown core and midtown where there are likely to be greater number of licensed premises and alcohol outlets, would result in higher alcohol expenditures. The spatial model was able to confirm that subway access had a significant positive association with both types of alcohol expenditures. Subway intercepts and lagged subway intercepts also proved to have a large effect on alcohol expenditures as it had the third largest effect after occupation and post-secondary education.

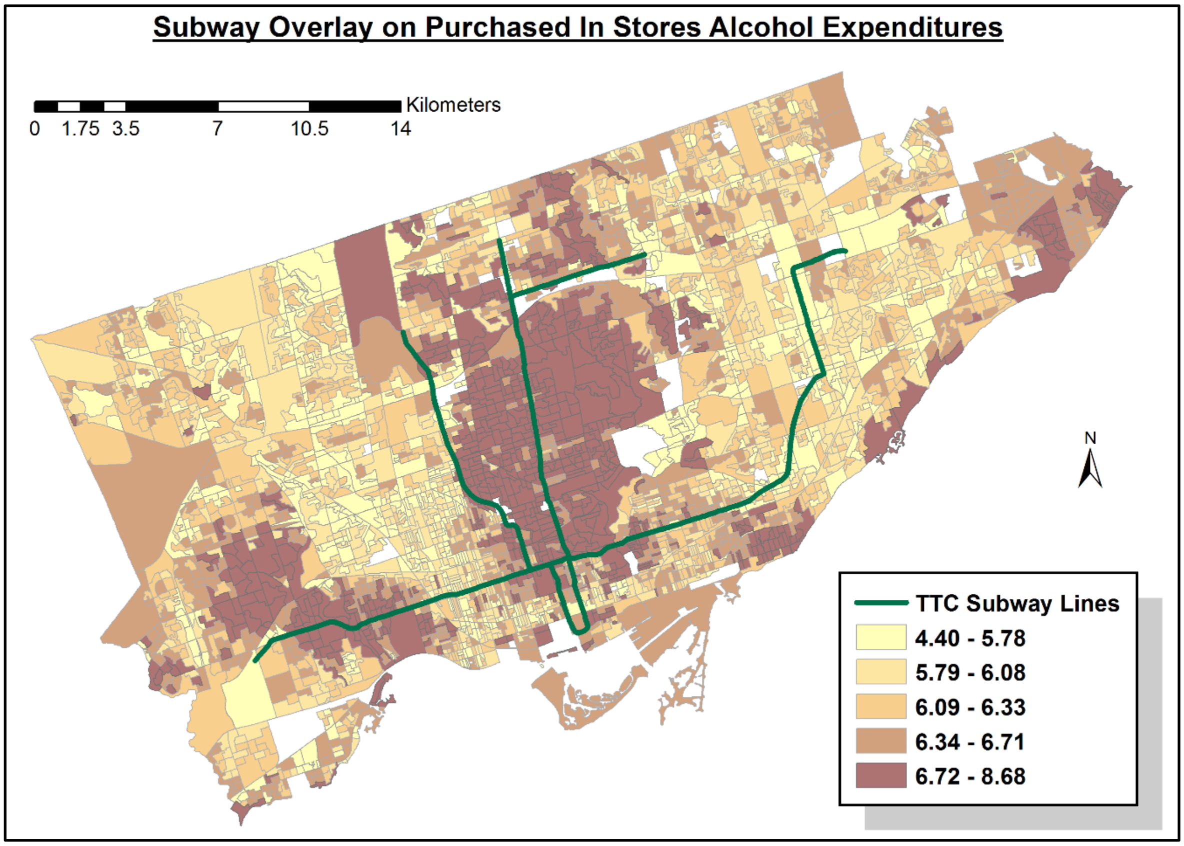



Figure 5.2.2.2: Map of Subway Overlay onto Purchased in Store Expenditures Quantile Map: The green line segments represents the subway lines in Toronto and have been overlaid on a quantile map of purchased in stores alcohol expenditures. The subway lines feature prominently in the major cluster of high alcohol expenditure in the middle of the city and in the cluster to the south-west of the city.

It should be noted that having a subway in a DA or in a neighbouring area could result in increased household expenditures in a variety of ways. It is possible that having subway access provides greater mobility, and therefore access to far away DAs with high alcohol outlet density. Alternatively, there might be pre-existing high alcohol outlet densities around areas with subway access in order to facilitate patrons. Finally, areas with subway access are generally much more expensive due to the ease of access to the public transportation system. A large proportion of people who reside in those areas may be those who have white-collar jobs, a post-secondary education and are more affluent. Future studies should try to further explore and elucidate this association.

Highways and lagged highway intercepts were not found to be significant predictors of either type of alcohol expenditure. It was hypothesized that the presence of a highway in the same DA or in a surrounding DA would lead to increased access to alcohol outlets and licensed premises, much like subway intercepts. A potential cause for this lack of association is that with DA areas being so small, having a stretch of highway in the area may actually decrease the amount of area for residences and licensed premises. Future studies should look at alternative ways to examine highway access. Possibilities include proximity to highway or presence of an on-ramp or off-ramp.

Alcohol outlet density and lagged alcohol outlet density pair were not significant predictors of purchased in stores expenditures. This was surprising given that an increased number of alcohol outlets in a Dissemination Area was hypothesized to increase the purchased in stores alcohol expenditures due to increased access to alcohol outlets. It is possible that due to the small size of the Dissemination Areas the amount of outlets in a Dissemination Area is not the driving factor; rather, it may be the distance to the nearest outlet. A rough measure of the straight-line distance (Euclidean distance) from the centroid of each DA to the closest outlet and was used as an explanatory variable. While this variable showed a positive significant association, the effect was extremely small. Future studies should utilize a more accurate measure of distance to examine the effects of proximity by using methods such as network analysis and buffer analysis.

Primary drinking restaurant density and lagged primary drinking restaurant density were found to have a positive association with licensed premises alcohol expenditure. This was the first study to quantify the relationship between the numbers of primary drinking restaurants in an area to the increase in licensed premises alcohol expenditure levels. It was hypothesized that increased primary restaurant density would result in increased expenditure due to greater access to alcohol.

Restaurant density and lagged restaurant density were also determined to be significantly positively associated with licensed premises alcohol expenditure levels. It was hypothesized that an increase in the density of restaurants would lead to an increase in alcohol expenditures, as there would be a greater availability of licensed premises. Restaurants covered all other licensed premises beyond those that just primarily served alcohol. Together with the primary drinking restaurant density variables, these two sets of variable confirm that increased access to licensed premises results in higher levels of alcohol expenditure.

## 5.3 Limitations and Strengths

### 5.3.1 Limitations

For small-area studies, one of the primary problems is the issue of the modifiable area unit problem (MAUP). The MAUP effect is the change in the association between explanatory variables and the outcome variable as a result of changes to the location and boundaries of areal units (22). Heywood formally defined the genesis of this problem as: "a problem arising from the imposition of artificial units of spatial reporting on continuous geographical phenomenon resulting in the generation of artificial spatial patterns" (126). As all geographical boundaries are arbitrary no matter how standardized they are, there is a possibility that MAUP can influence the results of any small-area level study. At best MAUP might be considered a nuisance, however, at worst, MAUP can lead to substantially altered results, bringing into question the reliability of the results (127).

The MAUP consists of two different components: 1) The zoning effect and 2) The scale effect (22, 128-130). The zoning effect (also known as the boundary effect) is when the number of area units remains constant; however, the locations of the boundaries change (22). Depending on where the boundaries are placed, the different zones can divide outcome and explanatory variables or concentrate them into certain areas, resulting in a different spatial structure for the study area. A hypothetical example of the zoning effect is if the 2011 census data contained then same number of Dissemination Areas for Toronto as the 2006 Census dataset; however, the borders were updated to follow new road networks. The second effect is the scale effect (also known as the aggregation effect) which occurs when the size of area units is varied, resulting in a differing number of spatial units for the study area (22). Using larger area units means combining or aggregating together area units from a smaller scale to create the larger area units. As a result of the aggregation of values to create the larger area unit, many local differences for a variable can be hidden as the large area unit provides the impression that the entire area is homogenous (129). An example of this is comparing the spatial structure of alcohol expenditure in Toronto at the Dissemination Area level to the spatial structure of alcohol expenditure in Toronto at the Census Tract level.

For this study, the Dissemination Area was chosen as the unit of analysis to balance data availability, limiting MAUP, finding local patterns and applicability to policy efforts. The Dissemination Area level was the smallest area level unit that contained census data for all of Canada and just as importantly, the smallest scale at which alcohol expenditure data was also available (99). The Dissemination Area level also followed six of the criteria for selecting the geographic scale for ecology studies, as defined by Arsenault et al.(131): 1) Widely accessible; 2) More likely to be homogenous in their outcome variable due to the small size of the DA; 3) Sufficient population size for stable results; 4) Consistency in population counts between each DA; 5) The size of each DA is also similar from one DA to the next; and 6) has a compact area. Compared with the commonly used Census Tracts, the Dissemination Area is far smaller and allows the identifying of much smaller local patterns of alcohol expenditure. The Dissemination Area also serves as a better approximation for a local neighbourhood due to its smaller areal size and population count, which is another fact in limiting the effects of MAUP. Finally, Census Tracts can be applied directly to land use and policy planning as the size of census tracts is similar to those of secondary planning areas and area-specific policy areas used by the City of Toronto (132). Dissemination Areas, however, can provide increased information on local patterns of alcohol expenditure while being easily associated with Census Tract areas to help inform policies for the City of Toronto.

A second major problem for small-area studies is the ecological fallacy. The ecological fallacy (cross-level bias) occurs when associations observed at a group level are used to make casual inferences at an individual-level (36, 132). This is problematic as an association between an outcome and explanatory variable at the group level may differ greatly compared to the actual casual association at the individual-level, leading to erroneous conclusions regarding individual-level associations (37, 133). An example from the context of this study is using the positive association found between having a post-secondary degree and alcohol expenditure at the Dissemination Area level to make an individual-level inference. It would be erroneous to assume that every individual with a post-secondary degree will spend more on alcohol. While this could be true for some people with post-secondary degrees, the real cause could be due to being exposed to a drink culture at post-secondary institutions or at workplaces because of their degree. For this study, it is important to remember that the unit of analysis is the Dissemination Area and that the findings are only applicable for understanding alcohol expenditure in Toronto at the Dissemination Area level.

While an ecological study should never be used on its own for casual inferences, there are a few ways in which to limit and offset the ecological bias as well as to overcome the bias. These methods are the following: the selection of the area unit, the use of instrumental variables and multilevel models. Aresenault et al.'s (131) criteria for selecting a geographical unit that limited ecological bias were: high intra-unit homogeneity, a consistent population size between areas, compactness in shape. Dissemination Areas fit this profile as they provide the smallest scale of geographical units that cover all of Canada, while having very compact shapes with small total area and as well have consistent population counts among all areas. The small area of the Dissemination Area and small population counts also help increase the chance of greater homogeneity among the individuals within the Dissemination Area. It is also possible to employ the use of instrumental variables (IV), originally an econometric idea to help overcome confounding by other unmeasured variables in ecological models (133, 134). In an ecological study, there is a possibility that an association found between an outcome Y and an explanatory X are confounded with the presence of the confounder, which shows up in the error term U. An instrumental variable is one that can predict X while being unrelated to the error term U and is also not directly related to Y (134, 135). An instrumental variable, which has these two properties, can be used as a proxy to measure the upper bound of the parameter estimate for the original X variable, with additional assumptions allowing the lower bound to be estimated as well. Loney and Nagelkerke (133) argue that in the case where an ecological variable acts as an IV for an individual-level association, it can be used to make causal inferences. Finally, the last method is employing a completely different model that combines both ecological and individual-level variables. The multilevel model can be used to examine both individual (compositional) and group-level (contextual) variables and determine the effect of both types of variables on the individual (22). In combining the different units of analysis, this overcomes both the scale dependency and ecological bias found in using ecological models (36). Future studies for alcohol expenditure should look to employ the multilevel model in order to examine the relationship between Dissemination Area level findings and the individual-level data. This will provide a greater ability for making causal inferences at the individual-level.

For this study 173 DAs were omitted due to missing socioeconomic data. In order to confirm whether there was cause for the missing data in those 173 Dissemination Areas, the Dissemination Area map with 3512 DAs was overlaid onto a Street Map made by OpenStreetMap (OSM), to examine what elements were contained within the missing DAs. The overlay in Figure 5.3.1.1 showed that the majority of these missing disseminations contained stretches of highways (401, DVP, Gardiner Expressway), roads (Markham Road, Lakeshore Blvd), green space (parks – i.e. Rouge Valley Park), railways and the Toronto Zoo.

**Overlay of DA Map onto Street Map**

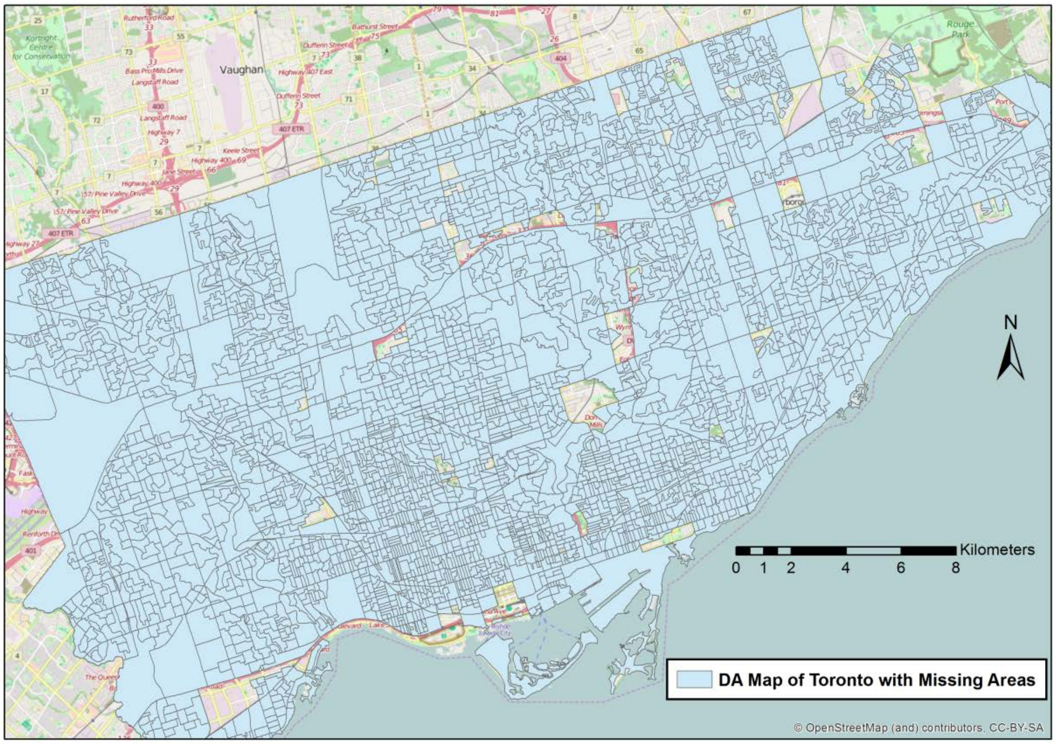



Figure 5.3.1.1: Dissemination Area Map Overlay on Street Map : The Dissemination Area Map was overlaid onto a Street Map made by OpenStreetMap (OSM). The holes from the missing DA's have the corresponding parts of the street map visible. The results indicate that the missing Dissemination Areas are due in large part to non-residential areas such as highways, parks and the Toronto Zoo.

A fourth major limitation for this study was the lack of alcohol pricing, taxation on alcohol and household alcohol consumption data. If the volume of alcohol could be controlled for, there would be a better understanding of how alcohol expenditure is associated with alcohol consumption, and therefore alcohol-related harms. As there was also no available pricing and tax data, the dependent variable of average household alcohol expenditure was assumed to implicitly capture the interaction between alcohol prices and demand for alcohol (32).

The fifth major limitation of this study was that there was no distinction between different types of alcoholic beverages in the expenditure data, as the alcohol expenditure data for all types of alcohol beverages was pooled together. If alcohol expenditures were delineated by type of alcoholic beverages, this study would be able to offer a better understanding of what type of alcoholic beverages drive alcohol purchases in the City of Toronto.

The sixth major limitation of the study was the inability to differentiate between licensed premise restaurants and non-licensed restaurants in the restaurant density variable. The primary drinking restaurants were restaurants that primarily focused on selling alcoholic beverages (i.e. pubs) (136) therefore, all of the primary drinking restaurants can be verified as licensed premises. Conversely, for the restaurant density variable, these were restaurants that focused on serving both food and alcohol (eating restaurants) (136); therefore, this variable constituted a mix of both restaurants that served alcohol and those that were not licensed to do so. The assumption was made that aside from fast food restaurants, the majority or at least a significant amount of the eating restaurants had a license to sell alcoholic beverages. As a result, ignoring the eating restaurants would be excluding a significant source of licensed premises. In terms of this study, it is possible then that the effect of restaurants was inflated. Future studies should look at re-running the same analysis after separating the licensed and non-licensed premises establishments in the eating restaurants dataset.

The seventh major limitation was the differences in years for the different datasets used to collect the alcohol expenditure data and the explanatory variables. In order to have the most complete set of socio-economic variables, the 2006 year of the Canadian Census was used as it was the last census year to have a mandatory long form of the census. None of the other three datasets had a 2006 year; therefore, for each dataset, data from the year closest to 2006 chosen. The 2010 Survey of Household Spending, 2008 DMTI EPOI data and 2010 DMTI CanMap Route Logistics datasets were selected. While there was unlikely to be dramatic changes for the alcohol expenditures and built-environment variables, the lag in time is still significant and should be addressed in future research. It should also be noted that with the advent of beer sales in grocery stores since December 2015 (137), future alcohol expenditure patterns may well change significantly due to this increase of retail alcohol outlet density. Future studies should also include these new outlets.

The eighth major limitation was the way in which built-environment variables were conceptualized. For subways and highways, both were measured binary variables that detected the presence of each built-environment in a Dissemination Area. Future studies may want to use measures that are even more specific for each, such as presence or proximity to a subway station and the presence or proximity to an on/off-ramp in a DA. For land use, the types of land use tested were not exhaustive by any means as only two land use categories were tested. Future studies may also want to examine possible connections between alcohol expenditures and other land use types at the Dissemination Area level. Lastly, this study was not able to include an effective measure of proximity to outlets. The straight-line distance (Euclidean) from the centroid (centre point) of each DA to the nearest alcohol outlet was calculated and included in a model for the purchased in stores expenditures. While the model showed that this nearest distance was significantly positively associated with alcohol expenditures, the effect was extremely miniscule. As this was an extremely crude measure of distance to outlets, this variable was removed from the final model. Future studies should look to include some sort of buffer analysis or network analysis to provide better insight into how distance to alcohol outlets affects alcohol expenditures.

The ninth major limitation concerns how alcohol expenditure data was obtained at the DA level. According to the data provider SimplyMap, their data vendor Environics Analytics

(EA) provided the household alcohol expenditures. EA utilized Statistics Canada's 2013 Survey of Household Spending in addition to demographic, household, and income data to estimate the alcohol expenditure values as part of their HouseholdSpend dataset. The HouseholdSpend dataset has both updated average expenditure estimates and total expenditure estimates for 275 different types of good and services purchased by Canadians. HouseholdSpend is produced annually with Dissemination Area disposable income data from DEP (Demographic Estimates and Projections) and PRIZM C2 segment spending patterns based on Statistics Canada's latest Survey of Household Spending. For the 2013 SHS, more than 24,000 private households across all provinces and territories were asked detailed questions on their spending patterns. HouseholdSpend is then distributed at standard levels of Census geography by EA using proprietary spatial interpolation methodologies. While EA claims that the rigorous methodologies were applied to create accurate and significant estimates, there is no way to verify that claim and to know the limitations of the methods they used. For the purposes of this study, the data was assumed to be accurate.

A tenth and final limitation is that the locations where alcohol is purchased and consumed do not necessarily coincide with where people reside. As the data is household expenditure data, the expenditures can come from various locations, including both outlets within a household's Dissemination Area, as well as expenditures in outlets residing in other Dissemination Areas. Future studies should look to acquire expenditure data that records locations of purchase in addition to household addresses. This data can be used to explore the relationship between spending on alcohol, place of consumption and proximity to outlets, leading to a better understanding of the catchment area for both licensed premises and purchased in stores outlets.

### 5.3.2 Strengths

In spite of these limitations, the current study carries plenty of strengths and adds some novel findings to the limited alcohol expenditure literature. This was the first study to determine and quantify the spatial structure of alcohol expenditure data at the small-area level in the City of Toronto. Furthermore, this study also applied local clustering techniques to identify specific areas with high levels of alcohol expenditure, which can serve as targets for alcohol-use reduction interventions. The use of the Dissemination Area unit as the geographical scale also increased the level of specificity, by allowing the Local Moran's I method to identify neighbourhoods that were hot-spots or cold-spots at a very fine spatial scale.

Another major strength of this study was the use of the small-area level as the unit of analysis. This was the first study to examine alcohol expenditure from an ecological perspective at the Dissemination Area level. Using an ecological perspective, this study was able to incorporate contextual variables into the study that operate above the individual-level. More specifically, this study examined built-environment correlates for alcohol expenditure. To our knowledge, this was the first study to examine how the physical environment in the form of built-environment variables affected alcohol expenditures.

Using spatial regression models, this study was able to control for a large spectrum of socio-economic and built-environment variables in addition to controlling for the clustering of alcohol expenditures and certain built-environment variables. As spatial regression models were able to account for geographic variation, this prevented the occurrence of Type II errors and biased parameter estimates .

Finally, as an entire package, the spatial data analysis methods and GIS tools employed by this study provided powerful tools to help map and visualize the alcohol expenditure data and explanatory variables. GIS tools also provided the ability to quick present the data at different spatial scales. These tools allowed the results of this study to be communicated clearly and effectively in the form of maps and other visualizations. As a result, the findings can be quickly disseminated and understood by readers of all levels of expertise.

## 5.4 Implications for Public Policy and Future Studies on Alcohol Expenditure

### 5.4.1 Implications for Public Policy

The first important finding from this study was that alcohol outlet density increased purchased in stores expenditures and similarly, primary drinking restaurant and restaurant density increased expenditures for licensed premises. This finding agrees with previous work by Gruenewald, Ponicki and Holder (93) that found that independent of alcoholic beverage pricing and while controlling for sales and availability, physical availability in the form of alcohol outlets increases the sales of alcoholic beverages. In the interest of reducing both chronic and acute alcohol-related harms due to greater consumption from greater sales of alcohol, alcohol outlet density and licensed restaurant density should be regulated. Previous research also supports the efficacy of regulating alcohol outlets as an effective public health tool for reducing alcohol sales and consumption. In their meta-analysis, Carla et al. propose a framework for reducing alcohol sales, harmful consumption patterns and alcohol-related harms. Figure 5.4.1.1 presents this framework.

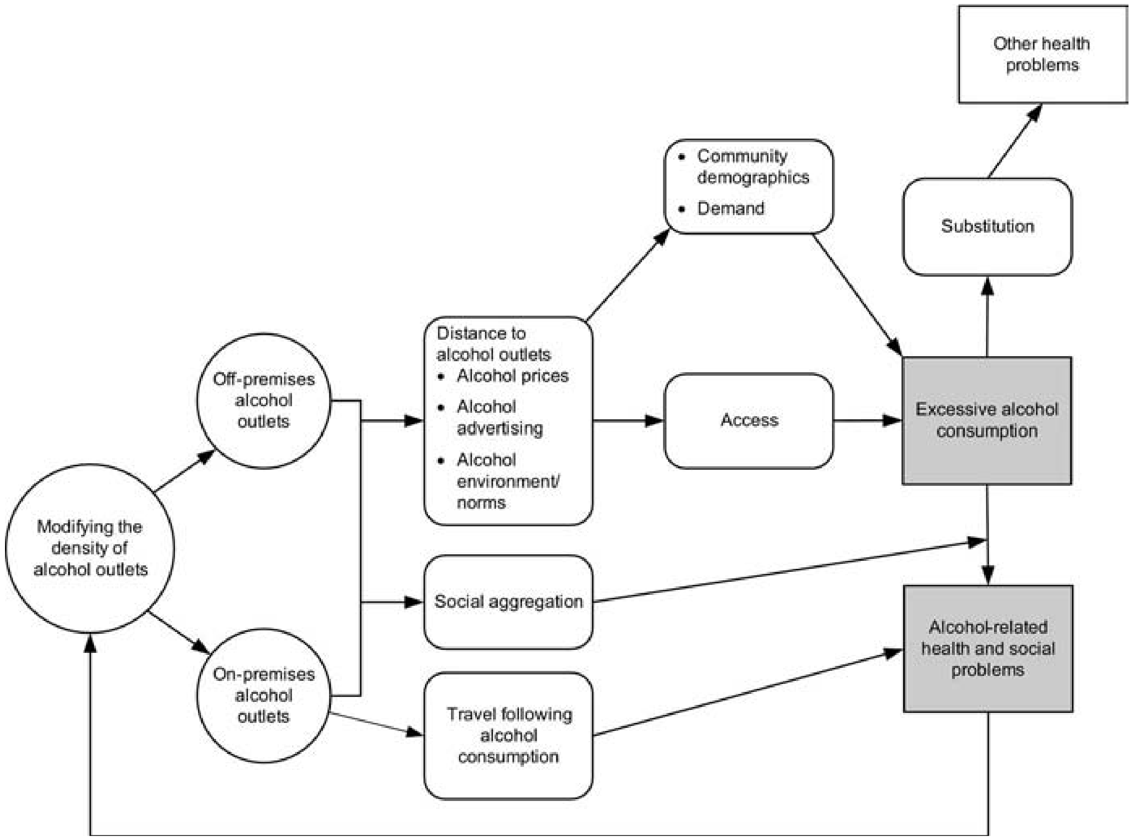



Figure 5.4.1.1 Framework for Alcohol outlet density regulation to reduce alcohol sales and harms (138)

Planners and policy makers can use GIS tools to determine the density of alcohol retailers in the neighbourhood before issuing an additional license to sell alcohol or before opening an additional alcohol outlet. Further to that, if a neighbourhood is high alcohol outlet density, decision makers can choose to reduce the amount of alcohol outlets or licensed premises in order to decrease alcohol availability. GIS methods can also be used to ensure that there is a certain amount of distance between alcohol outlets to ensure that they are properly spaced out and do not create high-density areas. In addition, the results of this study also showed that density of alcohol outlets in neighbouring areas also has a positive association with alcohol expenditures in an area. This implies that policy makers should also determine appropriate alcohol outlet densities for adjacent areas as well as for the immediate neighbourhood. Aside from limiting the number of alcohol outlet or restaurant licenses, policy makers can also change the proportion of those types of licences in target areas in order to alter drinking patterns (138). For example, there is evidence that bars and restaurants do not function identically despite both being on-premise outlets; therefore, changing the proportion of restaurants to bars could result in significant changes to alcohol expenditure patterns (59). GIS methods can also assist in ensuring that alcohol outlets are not located in close proximity to areas of concern such as schools. Areas where the population density is high can also be quickly identified and alcohol outlets assigned to the area can be restricted to account for the volume of people in those areas. The physical size of outlets can also be regulated in order to control the volume of people allowed in the facility, thereby limiting the overall amount of alcohol sales from the premises (138). Finally, other built-environment features such as the size of parking lots and the traffic surrounding alcohol outlets and restaurants can also be altered affect alcohol sales (138).

Hot-spots of alcohol outlet densities can also provide useful information to server training and certification programs such as Smart Serve in Ontario. Smart Serve could use the hot-spots by specifically monitoring them as high priority areas, to ensure that all servers working in those areas are properly trained and licensed. Furthermore, servers working in those areas could be further educated to be extra vigilant of over-serving alcohol to patrons.

Limiting alcohol outlet density can also serve to reduce alcohol-related harms to neighbourhoods with lower socio-economic status. Previous research showed that while the highest levels of consumption have been associated with neighbourhoods with higher socio-economic status, the highest alcohol outlet densities are often found in most deprived neighbourhoods, especially for off-site alcohol outlets (60, 139, 140). While, the effect is not fully understood, it is known these neighbourhoods suffer the most from acute alcohol-related harms from increased outlet density, such as assaultive violence, injury and other forms of criminality. Policymakers and urban planners can use overlay analysis to look for these deprived neighbourhoods which contain a high alcohol outlet density and enact measures to reduce the number of alcohol retailers in those neighbourhood.

The results of this study also have important law enforcement and healthcare planning implications. Studies have helped establish a positive association between alcohol-related harms and alcohol outlet density and restaurant density. These harms include violent crimes, injuries from accidents, motor-vehicle crashes, drunk driving, liver diseases, mortality and risky patterns of drinking such as binge drinking and underage drinking (90, 141-143). Law enforcement could use GIS tools as an efficient means of identifying high outlet density areas for policing. For health officials, emergency services can use GIS tools to map and monitor high alcohol outlet and restaurant densities for increased risk of alcohol-related accidents, injuries and fatalities. Health campaigns discouraging risky drinking behaviours and impaired driving can also be targeted to these areas to help mitigate the levels of alcohol-related harms.

The current study also identified the association between subways and alcohol expenditure as a key finding with policy implications. Previous studies have acknowledged the possibility that increased access by public transportation to areas with alcohol outlet densities could lead to greater alcohol expenditures (60). The findings in this study lend their support to this notion, as subways and having subway lines in neighbouring areas resulted in an increase in the level of alcohol expenditures. For policy makers, this information can serve as additional criteria by which to limit alcohol outlet densities. Areas that are proximal to subway access and subway stations could be regulated to have decreased alcohol outlet densities to help lower alcohol expenditures. However, this recommendation should also be balanced with other public health concerns such as the need to limit drunk driving, which may recommend alcohol outlets be situated closer to subways in order to reduce driving after drinking. Further research can be done to determine the appropriate densities for a range of distances around a subway station in order to balance both public health concerns.

Another significant finding was that both white-collar occupations and post-secondary degrees were positively associated with alcohol expenditures. For health planners, it would be important to first identify areas with an abundance of white-collar workplaces and then to regulate alcohol and drinking restaurants densities. For public health promotion, this demographic provides an important target population for possible alcohol-use reduction campaigns. It is possible that many drink establishments were made for white-collar clientele and have been effective in targeting this demographic. Public health could possibly make significant gains in harmful alcohol-use reduction by tailoring anti-harmful alcohol-use education campaigns for the white-collar demographic. The connection with post-secondary degrees also poses an interesting question for health researchers. It was hypothesized that higher educational attainment as indicated by post-secondary degrees led to more white-collar careers resulting in greater financial gain. However, while income had a positive effect on alcohol expenditures, the overall effect was very small. This indicates that there may be some underlying drinking culture associated with both or either explanatory variable. Previous research has identified the university context as influencing drinking behaviours (144). Future research should look for the causes of increased alcohol expenditure for post-secondary graduates and white-collar workers, as well as delineating specific white-collar occupations associated with high alcohol expenditure levels and include an examination of the post-secondary environment.

One final possible method regulatory method for alcohol outlets densities is to limit the hours of operations in areas which have been identified to have high alcohol outlet densities. Previous studies have found that regulating business hours has a significant impact on alcohol sales (145). As an example, policymakers can limit hours of operations on days such as holidays and weekends for certain alcohol outlets in high-density areas. GIS methods can also be employed to ensure that there is a certain amount distance between outlets that remain open.

Finally, the last major implication for this study is the use of GIS techniques and spatial methods. These tools can assist future public health research by helping with data visualization, examining data at different scales, accounting for spatial autocorrelation and identifying target areas for intervention. Using GIS methods different types of attribute data can be quickly linked to geographical points or areas and immediately visualized. This allows effective exploratory data analysis by allowing quick visual identification of possible spatial patterns in the data and secondly by allowing multiple variables to be overlaid onto the study region to see if there are possible associations between different variables of interest. Furthermore, GIS methods can be used to quickly change the unit of analysis in a study. For example, this study used the Dissemination Area as it was a very fine geographical unit which allowed this study to identify small local patterns for alcohol expenditures; however, policy may actually be enacted at a different scale such as the one hundred and forty historical Toronto neighbourhoods. This study transferred the alcohol clusters identified at the Dissemination Area to the Toronto neighbourhood level. Researchers can easily transfer the data to other policy relevant neighbourhood configurations such as Local Health Integration Networks (LHINs), public health units, Community Care Access Centre (CCAC) catchment areas and Ministry of Community and Social Services Regions. This study also demonstrated the application of four different spatial regression methods that can be used to study associations between any type of area level health outcome and explanatory variables while controlling for geographical clustering. Additionally, after identifying relevant factors, the cluster analysis used in this study can help identify specific hot-spots and cold-spots within the study region, where policies or programs can be enacted to help reduce harmful health outcomes.

### 5.4.2 Future Directions

This study was the first study to identify socio-economic and built-environment associations with alcohol expenditures at the Dissemination Area level. Although this study included numerous socio-economic and built-environment variables, the spatial error and spatial Durbin error models were better fits, indicating possible misspecifications due to missing variables. In particular, more built-environment variables should be explored and tested to help further understand the impact of the built-environment on alcohol expenditures. Both alcohol outlet densities by area and by population were tested in this study as there was no standardized way for calculating alcohol outlet density. Future studies can test other conceptualizations of alcohol outlet density such as density per roadway mile as proposed by Romley, Cohen, Ringel and Sturm (139). Other examples of possible built-environment variables that can be explored include a more comprehensive list of land uses, bus routes, density of roadways, hours of operations for outlets and more complete measures of proximity to alcohol outlets by using network analysis and buffer analysis.

Tobacco smoking is another modifiable behavioural risk factor that is often associated with alcohol use. The concurrent use of both substances has been well-document, as previous work has found that those who partake of one risk behaviour often take part in the other as well, making alcohol use and tobacco smoking close linked behaviours (146). Previous work by

Abdel-Ghany and Silver (15) has also identified tobacco expenditures as a significant predictor of alcohol expenditures. Future studies with access to such data should also account for tobacco expenditures as a possible predictor of alcohol expenditures.

The results of this study can be further extended by including individual-level factors for alcohol expenditure. Future studies can make use of multilevel models to examine factors at different ecological levels, namely the individual-level and the small-area level (Dissemination Areas) simultaneously, in order to quantify the proportion of the effect on alcohol expenditure from individual-level factors and small-area level factors. It is also important to note that most multilevel models are aspatial as they do not account for spatial autocorrelation in the model explicitly (22, 147). Therefore, many multilevel models use neighbourhood units without considering spatial autocorrelation between the area units as it assumes all spatial correlation can be reduced to within-neighbourhood correlation (147). Previous studies have shown that while multilevel models accounted for significant geographical variation, spatial autocorrelation unaccounted for by the multilevel still remained (148), implying that both within-neighbourhood and spatial processes can operate simultaneous (149). In order to account for spatial autocorrelation, a spatially-explicit multilevel model or a hierarchical geostatistical model should be used for the analysis.

To increase the robustness of results of this study, future studies can also make use of longitudinal data to examine alcohol expenditure patterns. The current study was a cross-sectional study that identified geographic variation and small-area level associations with alcohol expenditures for the year 2010. Using panel data with multiple years of alcohol expenditures can help strengthen the relationships identified and provide more evidence for a casual inference. As well, with the forthcoming addition of 450 grocery store licenses to sell beer and cider in Ontario

(150), this will likely have a significant impact on the spatial distribution of alcohol expenditures and alcohol use. Furthermore, with new public transit methods such as the Eglinton crosstown LRT slated to be completed over the next decade (151), these factors may also impact the spatial distribution of alcohol expenditures in a significant manner. Future studies can help visualize and identify changes in alcohol expenditure hot-spots and cold-spots over various years and examine what these novel factors will mean for alcohol expenditure levels overall.

# Chapter 6: Conclusion

A significant proportion of Canadians and Ontarians continue to consume alcohol above low-risk drinking guidelines despite the fact that harmful alcohol use is one of the main contributors to non-communicable diseases. Examining alcohol expenditures provides invaluable insights into the consumer demand for alcohol, the economic impact of alcohol expenditures and the proportion of income that households elect to spend on alcohol. Despite the importance of these findings, research on alcohol use has focused primarily on alcohol consumption, leaving a scarcity of alcohol expenditure studies. Furthermore, the alcohol expenditure work that does exist focuses on analysis at the individual-level, leaving contextual factors out of the analysis. The goals of this study were twofold: The first goal was to determine the geographical variation of alcohol expenditures in Toronto at the small-area level, and the second goal was to examine socio-economic and built-environment characteristics associated with alcohol expenditures at the small-area level while controlling for the geographical variation of alcohol expenditures.

The results of this study indicate that having a white-collar occupational status and higher educational attainment leads to much higher levels of alcohol expenditure at the Dissemination Area level. These results corroborated the findings in previous individual-level studies, and established small-area level associations between alcohol expenditure and socio-economic characteristics for the first time. Significant built-environment associations with alcohol expenditures were also identified for the first time in this study and underscore the importance of examining contextual factors as significant influences of health behaviours.

This study advances our knowledge and understanding of socio-economic and built-environment factors associated with alcohol expenditures at the Dissemination Area level for the

City of Toronto. Moreover, it adds considerably to the current understanding of alcohol expenditures by recognizing geographic influences on alcohol expenditure. Nevertheless, additional research is required to expand on this study's findings in order to further develop the overall understanding of alcohol expenditure. Specifically, future research should look to elucidate the specific white-collar occupations and post-secondary degrees associated with higher alcohol expenditure levels. Public health campaigns looking to reduce alcohol-related harms can target this demographic. The built-environment findings also provide significant implications for health policy, as policy makers and city planners can address harmful alcohol use by limiting restaurant and retail outlet densities. Furthermore, additional built-environment variables should be included in future analyses, as well as individual-level data. Spatially-explicit multilevel models should also be applied to help determine the relative effects of Dissemination Area level factors and individual-level factors on alcohol expenditures to provide an even more comprehensive understanding of factors influencing alcohol expenditures.

# Bibliography


1. Coreil J. Social and behavioral foundations of public health: SAGE Publications Inc; 2010.

2. (LDCP) LDCP. Addressing Alcohol Consumption and Alcohol-Related Harms at the Local Level: A Resource for Public Health Professionals in Ontario.

3. World Health Organization. Global status report on alcohol and health-2014: World Health Organization; 2014.

4. National Collaborating Centre for Mental Health, editor Alcohol Dependence and Harmful Alcohol Use2011: British Psychological Society.

5. World Health Organization. Harmful Use of Alcohol. In: health WHO-Ndam, editor. Online2009.

6. Shield KD, Taylor B, Kehoe T, Patra J, Rehm J. Mortality and potential years of life lost attributable to alcohol consumption in Canada in 2005. BMC Public Health. 2012;12(1):91.

7. Rehm J, Gnam W, Popova S, Baliunas D, Brochu S, Fischer B, et al. The costs of alcohol, illegal drugs, and tobacco in Canada, 2002. Journal of studies on alcohol and drugs. 2007;68(6):886.

8. Cancer Care Ontario Ontario Agency for Health Protection and Promotion (Public Health Ontario). Recommendations for a Healthier Ontario. Toronto: Queen’s Printer for Ontario: 2012.

9. Rehm J, Gmel G, Sempos CT, Trevisan M. Alcohol–related morbidity and mortality. Alcohol Res Health. 2003;140:C00-C97.

10. Frishman WH. Cardiac Disease. Alcohol: Science, Policy and Public Health. 2013:249.

11. Mukamal KJ. Vascular Disease. Alcohol: Science, Policy and Public Health. 2013:256.

12. Mukamal KJ, Chung H, Jenny NS, Kuller LH, Longstreth W, Mittleman MA, et al. Alcohol consumption and risk of coronary heart disease in older adults: the Cardiovascular Health Study. Journal of the American Geriatrics Society. 2006;54(1):30-7.

13. (CCSA) CCoSA. Canada's Low-Risk Alcohol Drinking Guidelines 2013 [cited 2015 April 26]. Available from: http://www.ccsa.ca/Resource%20Library/2012-Canada-Low-Risk-Alcohol-Drinking-Guidelines-Brochure-en.pdf.

14. (CAMH) CfAaMH. Canada's Low-Risk Alcohol Drinking Guidelines 2011 [cited 2015 April 26]. Available from: http://www.camh.ca/en/hospital/health_information/a_z_mental_health_and_addiction_information/alcohol/Pages/low_risk_drinking_guidelines.aspx.

15. Abdel-Ghany M, Silver JL. Economic and demographic determinants of Canadian households' use of and spending on alcohol. Family and Consumer Sciences Research Journal. 1998;27(1):62-90.

16. Grubesic TH, Pridemore WA, Williams DA, Philip-Tabb L. Alcohol outlet density and violence: the role of risky retailers and alcohol-related expenditures. Alcohol and alcoholism. 2013:agt055.

17. Nayga Jr RM, Capps Jr O. Analysis of alcohol consumption in the United States: Probability and level of intake. Journal of Food Distribution Research. 1994;25(2).

18. Blaylock JR, Blisard WN. Wine consumption by US men. Applied Economics. 1993;25(5):645-51.

19. Sharpe DL, Mohamed A-G, Kim H-Y, Hong G-S. Alcohol Consumption Decisions in Korea. Journal of Family and Economic Issues. 2001;22(1):7.

20. Gatrell AC, Elliott SJ. Geographies of health: an introduction: John Wiley & Sons; 2014.

21. Craglia M, Maheswaran R. GIS in public health practice: CRC press; 2004.

22. Cromley EK, McLafferty S. GIS and public health: Guilford Press; 2012.

23. Ward MD, Gleditsch KS. An introduction to spatial regression models in the social sciences. Manuscript at http://wwwfacultywashingtonedu/mdw 2007;8:2007.

24. Saelens BE, Handy SL. Built environment correlates of walking: a review. Medicine and science in sports and exercise. 2008;40(7 Suppl):S550.

25. McMichael AJ, Bolin B, Costanza R, Daily GC, Folke C, Lindahl-Kiessling K, et al. Globalization and the Sustainability of Human Health An ecological perspective. BioScience. 1999;49(3):205-10.

26. McMichael AJ. Prisoners of the proximate: loosening the constraints on epidemiology in an age of change. American journal of epidemiology. 1999;149(10):887-97.

27. Wieczorek WF, Hanson CE. New modeling methods. Alcohol Health Research World. 1997;21(4):331-9.

28. Auchincloss AH, Gebreab SY, Mair C, Roux AVD. A review of spatial methods in epidemiology, 2000–2010. Annual review of public health. 2012;33:107.

29. What We Do: Alcohol - Overview [press release]. 2015.

30. Boyle P, Boffetta P, Lowenfels AB, Burns H, Brawley O, Zatonski W, et al. Alcohol: Science, Policy and Public Health: Oxford University Press; 2013.

31. Livingston M, Callinan S. Underreporting in alcohol surveys: whose drinking is underestimated? Journal of studies on alcohol and drugs. 2015;76(1):158-64.

32. Tan A, Yen ST, Nayga Jr. RM. Factors Affecting Alcohol Purchase Decisions and Expenditures: A Sample Selection Analysis by Ethnicity in Malaysia. J Fam Econ Iss. 2009;30.

33. Yen ST, Jensen HH. Determinants of Household Expenditures on Alcohol. 1996.

34. Atkinson AB, Gomulka J, Stern NH. Spending on alcohol: evidence from the Family Expenditure Survey 1970-1983. The Economic Journal. 1990:808-27.

35. English D. Geographical epidemiology and ecological studies. Geographical and environmental epidemiology (P Elliott, J Cuzick, D English, R Stern, eds). 1992:1.

36. Rezaeian M, Dunn G, St Leger S, Appleby L. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. Journal of epidemiology and community health. 2007;61(2):98-102.

37. Richardson S. Statistical methods for geographical correlation studies. Oxford: University Press; 1992. p. 181-204.

38. Clarke KC, McLafferty SL, Tempalski BJ. On epidemiology and geographic information systems: a review and discussion of future directions. Emerging infectious diseases. 1996;2(2):85.

39. Bronfenbrenner U. Ecological models of human development. Readings on the development of children. 1994;2:37-43.

40. (VPA) WHO-VPA. The ecological framework 2015 [updated 2015; cited 2015 April 27].

41. Barton H. Land use planning and health and well-being. Land Use Policy. 2009;26:S115-S23.

42. Frank LD, Engelke PO. The built environment and human activity patterns: exploring the impacts of urban form on public health. Journal of Planning Literature. 2001;16(2):202-18.

43. Wang Y, Chau C, Ng W, Leung T. A review on the effects of physical built environment attributes on enhancing walking and cycling activity levels within residential neighborhoods. Cities. 2016;50:1-15.

44. Handy SL, Boarnet MG, Ewing R, Killingsworth RE. How the built environment affects physical activity: views from urban planning. American journal of preventive medicine. 2002;23(2):64-73.

45. Frank L, Kavage S, Devlin A. Health and the built environment: A review. Urban Design. 2011;4.

46. Kaczynski AT, Henderson KA. Environmental correlates of physical activity: a review of evidence about parks and recreation. Leisure Sciences. 2007;29(4):315-54.

47. Li F, Fisher KJ, Brownson RC, Bosworth M. Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults. Journal of epidemiology and community health. 2005;59(7):558-64.

48. Saarloos D, Kim J-E, Timmermans H. The built environment and health: Introducing individual space-time behavior. International Journal of Environmental Research and Public Health. 2009;6(6):1724-43.

49. Perdue WC, Stone LA, Gostin LO. The built environment and its relationship to the public's health: the legal framework. American Journal of Public Health. 2003;93(9):1390-4.

50. Jackson RJ. The impact of the built environment on health: an emerging field. American Journal of Public Health. 2003;93(9):1382-4.

51. Canada S. Heavy drinking, 2013 Ottawa2013 [updated July 12, 2014; cited 2015 April 25]. Available from: http://www.statcan.gc.ca/pub/82-625-x/2014001/article/14019-eng.htm#n1.

52. Summary Tables - Heavy drinking, by age group and sex (Number of persons) [Internet]. 2013 [cited 2015-04-25]. Available from: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/health79a-eng.htm.

53. Canada H. Table 9: Prevalence of alcohol use and exceeding LRDG, by age, CAS 2004, CADUMS 2008 - 2011. 2011.

54. Wilsnack RW, Wilsnack SC, Kristjanson AF, Vogeltanz-Holm ND, Gmel G. Gender and alcohol consumption: patterns from the multinational GENACIS project. Addiction. 2009;104(9):1487-500.

55. Gius MP. An estimate of the effects of age, taxes, and other socioeconomic variables on the alcoholic beverage demand of young adults. The Social Science Journal. 2005;42(1):13-24.

56. Blaylock JR, Blisard WN. Women and the demand for alcohol: Estimating Participation and Consumption. The Journal of Consumer Affairs. 1993;27(2).

57. Berry JG, Pidd K, Roche AM, Harrison JE. Prevalence and patterns of alcohol use in the Australian workforce: findings from the 2001 National Drug Strategy Household Survey. Addiction. 2007;102(9):1399-410.

58. Pan S, Fang C, Malaga J. Alcoholic beverage consumption in China: a censored demand system approach. Applied Economics Letters. 2006;13(15):975-9.

59. Lipton R, Yang X, A. Braga A, Goldstick J, Newton M, Rura M. The geography of violence, alcohol outlets, and drug arrests in Boston. American journal of public health. 2013;103(4):657-64.

60. Pollack CE, Cubbin C, Ahn D, Winkleby M. Neighbourhood deprivation and alcohol consumption: does the availability of alcohol play a role? International journal of epidemiology. 2005;34(4):772-80.

61. Townsend P. Deprivation. Journal of social policy. 1987;16(02):125-46.

62. Pampalon R, Hamel D, Gamache P, Raymond G. A deprivation index for health planning in Canada. Chronic Dis Can. 2009;29(4):178-91.

63. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. Journal of epidemiology and community health. 2001;55(2):111-22.

64. Karvonen S, Rimpelä AH. Urban small area variation in adolescents' health behaviour. Social Science & Medicine. 1997;45(7):1089-98.

65. Ecob R, Macintyre S. Small area variations in health related behaviours; do these depend on the behaviour itself, its measurement, or on personal characteristics? Health & place. 2000;6(4):261-74.

66. Rice N, Carr-Hill R, Dixon P, Sutton M. The influence of households on drinking behaviour: a multilevel analysis. Social Science & Medicine. 1998;46(8):971-9.

67. Scribner RA, Cohen DA, Fisher W. Evidence of a structural effect for alcohol outlet density: a multilevel analysis. Alcoholism: Clinical and Experimental Research. 2000;24(2):188-95.

68. Milan A, Bohnert N. Portrait of families and living arrangements in Canada. Families, households and marital status. 2011:1-20.

69. Census snapshot of Canada — Families [Internet]. Government of Canada. 2006 [cited 2015-04-25]. Available from: http://www.statcan.gc.ca/pub/11-008-x/2007006/article/10380-eng.htm.

70. Duncan GJ, Brooks-Gunn J, Klebanov PK. Economic deprivation and early childhood development. Child development. 1994;65(2):296-318.

71. Chzhen Y, Bradshaw J. Lone parents, poverty and policy in the European Union. Journal of European Social Policy. 2012;22(5):487-506.

72. Ramful P, Zhao X. Individual Heterogeneity in Alcohol Consumption: The Case of Beer, Wine and Spirits in Australia*. Economic Record. 2008;84(265):207-22.

73. Harris MN, Ramful P, Zhao X. An ordered generalised extreme value model with application to alcohol consumption in Australia. Journal of health economics. 2006;25(4):782-801.

74. Canada S. Aboriginal Peoples in Canada: First Nations People, Métis and Inuit Ottawa2014 [updated 2014-03-28; cited 2015 04-25]. Available from: http://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-011-x/99-011-x2011001-eng.cfm.

75. The Health of Inuit, Métis and First Nations Adults Living Off-Reserve in Canada: The Impact of Socio-economic Status on Inequalities in Health [Internet]. 2010 [cited 2015-04-25]. Available from: http://www.statcan.gc.ca/pub/82-622-x/2010004/findings-resultats-eng.htm.

76. Elton-Marshall T, Leatherdale ST, Burkhalter R. Tobacco, alcohol and illicit drug use among Aboriginal youth living off-reserve: results from the Youth Smoking Survey. Canadian Medical Association Journal. 2011;183(8):E480-E6.

77. Cancer Care Ontario OAfHPaPPHO. Cancer Risk Factors in Ontario: Alcohol. Toronto: Queen's Printer for Ontario: 2014.

78. DMTI Spatial Inc, cartographer CanMap RouteLogistics Land Use. Markham, Ontario: DMTI Spatial Inc.; 2010.

79. Grubesic TH, Murray AT, Pridemore WA, Tabb LP, Liu Y, Wei R. Alcohol beverage control, privatization and the geographic distribution of alcohol outlets. BMC public health. 2012;12(1):1015.

80. Snowden AJ, Pridemore WA. Alcohol Outlets, Social Disorganization, Land Use, and Violence in a Large College Town Direct and Moderating Effects. Criminal Justice Review. 2013;38(1):29-49.

81. Pridemore WA, Grubesic TH. A spatial analysis of the moderating effects of land use on the association between alcohol outlet density and violence in urban areas. Drug and alcohol review. 2012;31(4):385-93.

82. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. Public health reports. 2001;116(5):404.

83. Ashe M, Jernigan D, Kline R, Galaz R. Land use planning and the control of alcohol, tobacco, firearms, and fast food restaurants. American Journal of Public Health. 2003;93(9):1404-8.

84. Curtin K, Pawloski L. Methodological Prospects with the Use of Geographic Information Systems and Spatial Statistics for the Evaluation of Alcohol Policy. World Medical & Health Policy. 2014;6(3):242-61.

85. Gruenewald PJ, Remer L, Lipton R. Evaluating the alcohol environment: community geography and alcohol problems. Alcohol Research and Health. 2002;26(1):42-8.

86. Toomey TL, Erickson DJ, Carlin BP, Lenk KM, Quick HS, Jones AM, et al. The association between density of alcohol establishments and violent crime within urban neighborhoods. Alcoholism: Clinical and Experimental Research. 2012;36(8):1468-73.

87. Zhu L, Gorman DM, Horel S. Alcohol outlet density and violence: a geospatial analysis. Alcohol and alcoholism. 2004;39(4):369-75.

88. Britt HR, Carlin BP, Toomey TL, Wagenaar AC. Neighborhood level spatial analysis of the relationship between alcohol outlet density and criminal violence. Environmental and Ecological Statistics. 2005;12(4):411-26.

89. Gruenewald PJ, Remer L. Changes in outlet densities affect violence rates. Alcoholism: Clinical and Experimental Research. 2006;30(7):1184-93.

90. Livingston M. Alcohol outlet density and assault: a spatial analysis. Addiction. 2008;103(4):619-28.

91. Gruenewald PJ. Regulating availability: how access to alcohol affects drinking and problems in youth and adults. Alcohol research and health. 2011;34(2):248.

92. Norström T. Outlet density and criminal violence in Norway, 1960-1995. Journal of studies on alcohol. 2000;61(6):907-11.

93. Gruenewald PJ, Ponicki WR, Holder HD. The Relationship of Outlet Densities to Alcohol Consumption: A Time Series Cross-Sectional Analysis. Alcoholism: Clinical and Experimental Research. 1993;17(1):38-47.

94. Bluthenthal RN, Cohen DA, Farley TA, Scribner R, Beighley C, Schonlau M, et al. Alcohol availability and neighborhood characteristics in Los Angeles, California and southern Louisiana. Journal of Urban Health. 2008;85(2):191-205.

95. Popova S, Giesbrecht N, Bekmuradov D, Patra J. Hours and days of sale and density of alcohol outlets: impacts on alcohol consumption and damage: a systematic review. Alcohol and Alcoholism. 2009;44(5):500-16.

96. Theall KP, Scribner R, Cohen D, Bluthenthal RN, Schonlau M, Lynch S, et al. The neighborhood alcohol environment and alcohol-related morbidity. Alcohol and alcoholism. 2009;44(5):491-9.

97. Canadian Alcohol and Drug Use Monitoring Survey (CADMUS). Table 8: Prevalence of alcohol use and exceeding LRDG, total population, CAS 2004, CADUMS 2008 - 2011. Health Canada; 2011.

98. (CADMUS) CAaDUMS. Table 7: Main 2012 CADUMS indicators by province - Alcohol. Health Canada; 2012.

99. Canada S. Dissemination block (DB) Ottawa2011 [updated 2012-12-04; cited 2015 April 26]. Available from: http://www12.statcan.gc.ca/census-recensement/2011/ref/dict/geo014-eng.cfm.

100. Statistics Canada. Census metropolitan area (CMA) and census agglomeration (CA) 2009 [updated December 11, 2007; cited 2016 April 4, 2016]. Available from: http://www12.statcan.ca/census-recensement/2006/ref/dict/geo009-eng.cfm.

101. City of Toronto. Neighbourhood Profiles 2016. Available from: http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=ae17962c8c3f0410VgnVCM10000071d60f89RCRD.

102. City of Toronto. CREATION OF TORONTO "NEIGHBOURHOOD" BOUNDARIES. 2016.

103. Social Development Finance & Administration, cartographer Neighbourhoods (UTM6) [Shapefile]2014.

104. Statistics Canada. Survey of Household Spending (SHS) - Detailed information for 2013: Statistics Canada; 2013 [updated 2015-01-21; cited 2015 August 7, 2015]. Available from: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3508.

105. Statistics Canada. 2006 Census - Taking a Census of Population 2009 [updated 2009-11-20; cited 2015 Friday August 7]. Available from: http://www12.statcan.gc.ca/census-recensement/2006/ref/dict/overview-apercu/pop2-eng.cfm#27.

106. Statistics Canada. Relationship to household reference person (Person 1): Statistics Canada; 2012 [updated 2012-12-04; cited 2015 August 7, 2015]. Available from: http://www12.statcan.gc.ca/census-recensement/2011/ref/dict/pop117-eng.cfm.

107. University of Waterloo Geospatial Centre. DMTI CanMap enhanced points of interest (EPOI) 2013 [updated 2013; cited 2015 Friday August 7]. Available from: https://uwaterloo.ca/library/geospatial/collections/canadian-geospatial-data-resources/dmti-canmap-enhanced-points-interest-epoi.

108. University of Waterloo Geospatial Centre. DMTI CanMap route logistics 2014 [updated 2014; cited 2015 Friday August 7]. Available from: https://uwaterloo.ca/library/geospatial/collections/canadian-geospatial-data-resources/canada/dmti-canmap-route-logistics.

109. City of Toronto. Open Data - Data Catalogue: Neighbourhoods 2016 [updated June 6, 2014; cited 2016 June 20, 2016]. Available from: http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=04b489fe9c18b210VgnVCM1000003dd60f89RCRD.

110. ESRI. How Spatial Autocorrelation (Global Moran's I) works: ESRI; 2012 [updated 3/7/2012; cited 2015 Friday August 7]. Available from: http://help.arcgis.com/EN/arcgisdesktop/10.0/help/index.html#//005p0000000t000000.

111. Sawada M. Global Spatial Autocorrelation indices–Moran's I, Geary's C and the General Cross-Product Statistic. Laboratory of Paleoclimatology and Climatology, Dept Geography, University of Ottawa,(Mimeo). 2001.

112. LeSage JP. The theory and practice of spatial econometrics. University of Toledo Toledo, Ohio. 1999;28:33.

113. Viton PA. Notes on spatial econometric models. City and regional planning. 2010;870(03):9-10.

114. Paradis E. Moran's Autocorrelation Coefficient in Comparative Methods. R Vignette. 2014.

115. Anselin L. Local indicators of spatial association—LISA. Geographical analysis. 1995;27(2):93-115.

116. GIS Geography. Spatial Autocorrelation and Moran's I in GIS 2015 [cited 2016 April 4, 2016]. Available from: http://gisgeography.com/spatial-autocorrelation-moran-i-gis/.

117. LeSage JP. What regional scientists need to know about spatial econometrics. Available at SSRN 2420725. 2014.

118. Anselin L. Spatial econometrics. A companion to theoretical econometrics. 2001;310330.

119. LeSage JP. An introduction to spatial econometrics. Revue d'économie industrielle. 2008(123):19-44.

120. Elhorst JP. Applied spatial econometrics: raising the bar. Spatial Economic Analysis. 2010;5(1):9-28.

121. Anselin L. Exploring spatial data with GeoDaTM: a workbook. Urbana. 2004;51:61801.

122. Xu Y, Kennedy E. An Introduction to Spatial Analysis in Social Science Research.

123. Zhukov YM. Applied spatial statistics in R, Section 6. Geostatistics[Online] Available: http://www people fas harvard edu/~ zhukov/Spatial5 pdf. 2010.

124. Kephart C. Interpret Regression Coefficient Estimates - {level-level, log-level, level-log & log-log regression} [cited 2016 April 2, 2016]. Available from: http://www.cazaar.com/ta/econ113/interpreting-beta.

125. Nayga RM. Sample selectivity models for away from home expenditures on wine and beer. Applied Economics. 1996;28(11):1421-5.

126. Heywood I, Cornelius S, Carver S. An introduction to Geographical Information Systems. 1998.

127. Openshaw S, Openshaw S, editors. The modifiable areal unit problem1984: Geo Abstracts University of East Anglia.

128. Parenteau M-P, Sawada MC. The modifiable areal unit problem (MAUP) in the relationship between exposure to NO2 and respiratory health. International journal of health geographics. 2011;10(1):1-15.

129. Oliver L. Shifting boundaries, shifting results: The modifiable areal unit problem. Available at: httpzllwww geog ubc ca/courses/geog516/talks_2001/scale_maup html Accessed January. 2001;27:2005.

130. ReStore @ Nation Centre for Research Methods. The Modifiable Areal Unit Problem [cited 2016 April 27, 2016]. Available from: http://www.restore.ac.uk/geo-refer/91023cwors00y00000000.php.

131. Arsenault J, Michel P, Berke O, Ravel A, Gosselin P. How to choose geographical units in ecological studies: Proposal and application to campylobacteriosis. Spatial and spatio-temporal epidemiology. 2013;7:11-24.

132. Quick M. Exploring crime in Toronto, Ontario with applications for law enforcement planning: Geographic analysis of hot spots and risk factors for expressive and acquisitive crimes: University of Waterloo; 2013.

133. Loney T, Nagelkerke NJ. The individualistic fallacy, ecological studies and instrumental variables: a causal interpretation. Emerging themes in epidemiology. 2014;11(1):1.

134. Greenland S. An introduction to instrumental variables for epidemiologists. International journal of epidemiology. 2000;29(4):722-9.

135. Bond S. instrumental variables1. In: Bond S, editor. First Year Econometrics: Nuffield College,

University of Oxford,; 2014.

136. DMTI Spatial Inc. SIC Descriptions. 2010.

137. CBC News. Kathleen Wynne heralds start of beer sales at Ontario grocery stores. CBCNEWS2015.

138. Campbell CA, Hahn RA, Elder R, Brewer R, Chattopadhyay S, Fielding J, et al. The effectiveness of limiting alcohol outlet density as a means of reducing excessive alcohol consumption and alcohol-related harms. American journal of preventive medicine. 2009;37(6):556-69.

139. Romley JA, Cohen D, Ringel J, Sturm R. Alcohol and environmental justice: the density of liquor stores and bars in urban neighborhoods in the United States. Journal of studies on alcohol and drugs. 2007;68(1):48-55.

140. Hay GC, Whigham PA, Kypri K, Langley JD. Neighbourhood deprivation and access to alcohol outlets: a national study. Health & place. 2009;15(4):1086-93.

141. Stockwell T, Zhao J, Macdonald S, Vallance K, Gruenewald P, Ponicki W, et al. Impact on alcohol-related mortality of a rapid rise in the density of private liquor outlets in British Columbia: a local area multi-level analysis. Addiction. 2011;106(4):768-76.

142. Pridemore WA, Grubesic TH. Alcohol Outlets and Community Levelsof Interpersonal Violence: Spatial Density, Outlet Type, and Seriousness of Assault. Journal of Research in Crime and Delinquency. 2011:0022427810397952.

143. Gruenewald PJ, Johnson FW, Treno AJ. Outlets, drinking and driving: a multilevel analysis of availability. Journal of studies on alcohol. 2002;63(4):460-8.

144. Weitzman ER, Folkman A, Folkman MKL, Wechsler H. The relationship of alcohol outlet density to heavy and frequent drinking and drinking-related problems among college students at eight universities. Health & place. 2003;9(1):1-6.

145. Hahn RA, Kuzara JL, Elder R, Brewer R, Chattopadhyay S, Fielding J, et al. Effectiveness of policies restricting hours of alcohol sales in preventing excessive alcohol consumption and related harms. American journal of preventive medicine. 2010;39(6):590-604.

146. Drobes DJ. Concurrent alcohol and tobacco dependence: mechanisms and treatment. Alcohol Research and Health. 2002;26(2):136-42.

147. Chaix B, Merlo J, Subramanian S, Lynch J, Chauvin P. Comparison of a spatial perspective with the multilevel analytical approach in neighborhood studies: the case of mental and behavioral disorders due to psychoactive substance use in Malmö, Sweden, 2001. American journal of epidemiology. 2005;162(2):171-82.

148. Chaix B, Merlo J, Chauvin P. Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France. Journal of Epidemiology and Community Health. 2005;59(6):517-26.

149. Arcaya M, Brewster M, Zigler CM, Subramanian S. Area variations in health: A spatial multilevel modeling approach. Health & place. 2012;18(4):824-31.

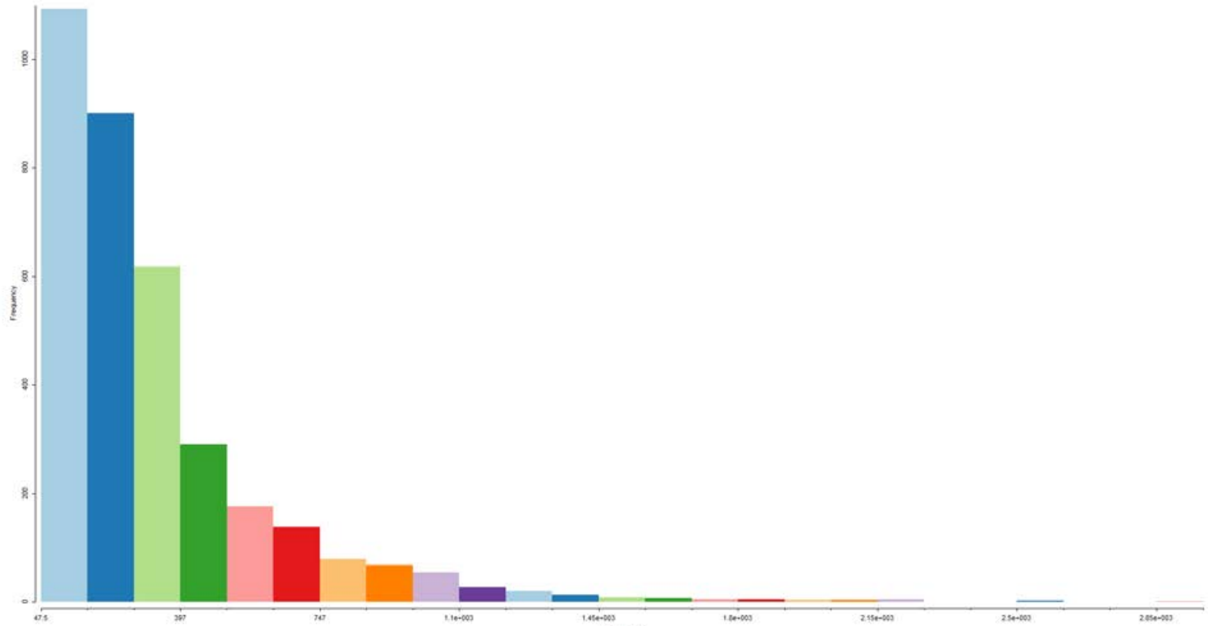
150. Ministry of Finance (Ontario). Beer sales in grocery stores Online2016 [updated April 18, 2016; cited 2016 April 27]. Available from: https://www.ontario.ca/page/beer-sales-grocery-stores.

151. Metrolinx. FAQs 2016 [cited 2016 April 27]. Available from: http://www.thecrosstown.ca/the-project/faq.

# Appendix

## A1. Histograms

Histogram for Untransformed Licensed Premises Expenditure Rate

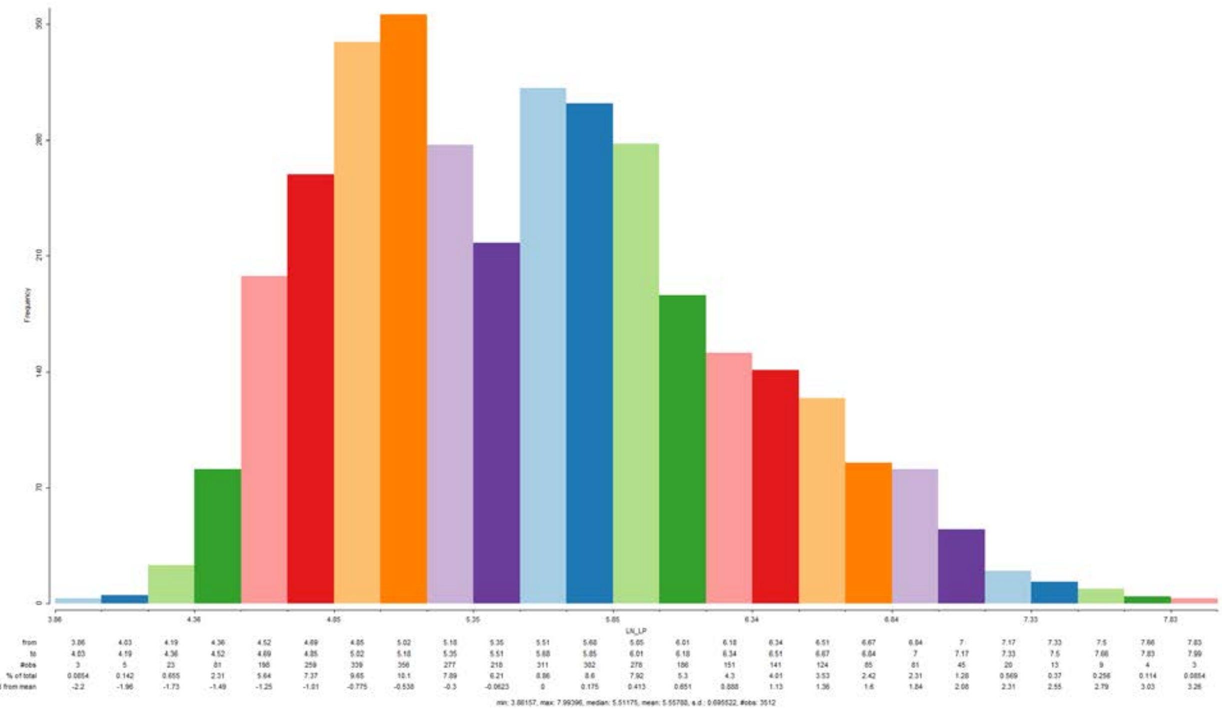



| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| from | 47.5 | 164 | 281 | 397 | 514 | 631 | 747 | 864 | 980 | 1.1e+003 | 1.21e+003 | 1.33e+003 | 1.45e+003 |
| to | 164 | 281 | 397 | 514 | 631 | 747 | 864 | 980 | 1.1e+003 | 1.21e+003 | 1.33e+003 | 1.45e+003 | 1.56e+003 |
| #obs | 1092 | 901 | 618 | 291 | 176 | 138 | 78 | 68 | 55 | 27 | 20 | 13 | 8 |
| % of total | 31.1 | 25.7 | 17.6 | 8.29 | 5.01 | 3.93 | 2.22 | 1.94 | 1.57 | 0.769 | 0.569 | 0.37 | 0.228 |
| sd from mean | -0.596 | -0.195 | 0 | 0.206 | 0.607 | 1.01 | 1.41 | 1.81 | 2.21 | 2.61 | 3.01 | 3.41 | 3.81 |

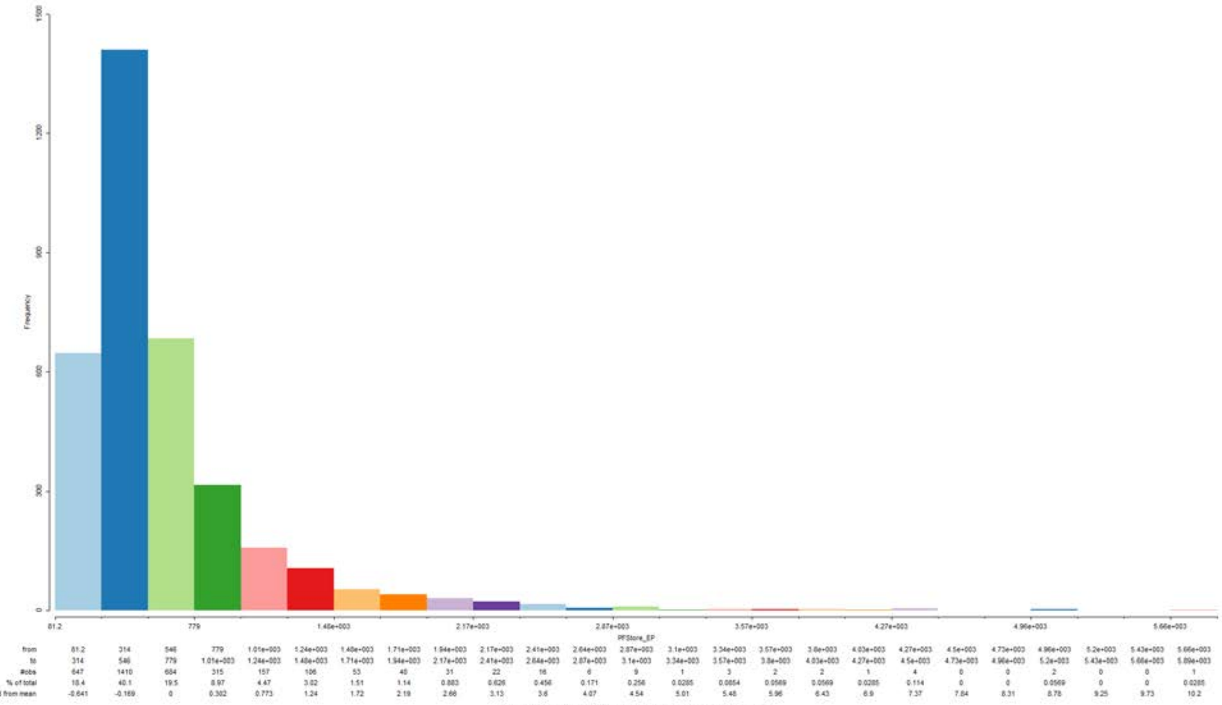
| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| from | 1.56e+003 | 1.68e+003 | 1.8e+003 | 1.91e+003 | 2.03e+003 | 2.15e+003 | 2.26e+003 | 2.38e+003 | 2.5e+003 | 2.61e+003 | 2.73e+003 | 2.85e+003 |
| to | 1.68e+003 | 1.8e+003 | 1.91e+003 | 2.03e+003 | 2.15e+003 | 2.26e+003 | 2.38e+003 | 2.5e+003 | 2.61e+003 | 2.73e+003 | 2.85e+003 | 2.96e+003 |
| #obs | 6 | 4 | 4 | 3 | 3 | 4 | 0 | 0 | 2 | 0 | 0 | 1 |
| % of total | 0.171 | 0.114 | 0.114 | 0.0854 | 0.0854 | 0.114 | 0 | 0 | 0.0569 | 0 | 0 | 0.0285 |
| sd from mean | 4.21 | 4.61 | 5.01 | 5.42 | 5.82 | 6.22 | 6.62 | 7.02 | 7.42 | 7.82 | 8.22 | 8.62 |

min: 47.54, max: 2963.02, median: 247.585, mean: 337.511, s.d.: 291.009, #obs: 3512

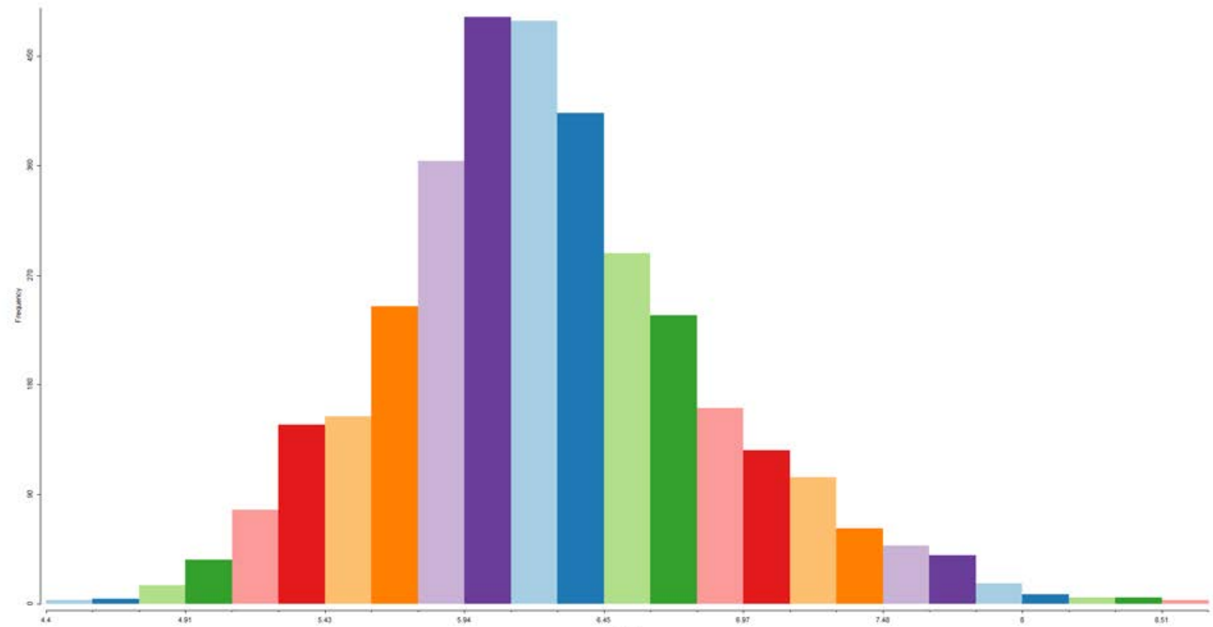
Histogram for Log-transformed Licensed Premises Expenditure Rate



Histogram for Untransformed Purchased in Stores Expenditure Rate



Histogram for Log-transformed Purchased in Stores Expenditure Rate




| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| from | 4.4 | 4.57 | 4.74 | 4.91 | 5.08 | 5.25 | 5.43 | 5.6 | 5.77 | 5.94 | 6.11 | 6.28 | 6.45 |
| to | 4.57 | 4.74 | 4.91 | 5.08 | 5.25 | 5.43 | 5.6 | 5.77 | 5.94 | 6.11 | 6.28 | 6.45 | 6.62 |
| #obs | 3 | 4 | 15 | 36 | 77 | 147 | 154 | 244 | 364 | 482 | 479 | 403 | 288 |
| % of total | 0.0854 | 0.114 | 0.427 | 1.03 | 2.19 | 4.19 | 4.38 | 6.95 | 10.4 | 13.7 | 13.6 | 11.5 | 8.2 |
| sd from mean | -2.79 | -2.5 | -2.22 | -1.93 | -1.65 | -1.36 | -1.08 | -0.794 | -0.509 | -0.224 | 0 | 0.0606 | 0.346 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| from | 6.62 | 6.8 | 6.97 | 7.14 | 7.31 | 7.48 | 7.65 | 7.82 | 8 | 8.17 | 8.34 | 8.51 |
| to | 6.8 | 6.97 | 7.14 | 7.31 | 7.48 | 7.65 | 7.82 | 8 | 8.17 | 8.34 | 8.51 | 8.68 |
| #obs | 237 | 161 | 126 | 104 | 62 | 48 | 40 | 17 | 8 | 5 | 5 | 3 |
| % of total | 6.75 | 4.58 | 3.59 | 2.96 | 1.77 | 1.37 | 1.14 | 0.484 | 0.228 | 0.142 | 0.142 | 0.0854 |
| sd from mean | 0.631 | 0.916 | 1.2 | 1.49 | 1.77 | 2.06 | 2.34 | 2.63 | 2.91 | 3.2 | 3.48 | 3.77 |

min: 4.39667, max: 8.68171, median: 6.19045, mean: 6.24565, s.d.: 0.601462, #obs: 3512

# A2. Bivariate Spatial Error Regressions for Licensed Premises Expenditures

**Bivariate Regressions for Licensed Premises Expenditures - 1**

| | ***Dependent variable:*** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **licensed premises expenditure** | | | | | | | | | | | |
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** | **(10)** | **(11)** | **(12)** |
| **Socio-Economic Variables** | | | | | | | | | | | | |
| Percentage of 20 to 44 Year Olds | -0.510*** | | | | | | | | | | | |
| Male to Female Ratio | | 0.037** | | | | | | | | | | |
| Percentage White | | | 0.010*** | | | | | | | | | |
| Percentage of Black Ethnicity | | | | -0.031*** | | | | | | | | |
| Percentage of Chinese Ethnicity | | | | | -0.002* | | | | | | | |
| Percentage of Filipino Ethnicity | | | | | | -0.005** | | | | | | |
| Percentage of South Asian Ethnicity | | | | | | | -0.029*** | | | | | |
| Percentage of Persons in White-collar Occupations | | | | | | | | 1.518*** | | | | |
| Percentage of Persons in Blue-collar Occupations | | | | | | | | | -1.586*** | | | |
| Median Household Income - After-Tax | | | | | | | | | | 0.00001*** | | |
| Percentage of Income from | | | | | | | | | | | -0.035*** | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Transfer Payments | | | | | | | | | | | | |
| Percentage of Apartments - 5 floors or more | | | | | | | | | | | | -0.005*** |
| Constant | 5.732*** | 5.504*** | 4.960*** | 5.551*** | 5.561*** | 5.555*** | 5.551*** | 4.747*** | 5.814*** | 4.906*** | 5.942*** | 5.637*** |
| **Observations** | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 |
| **Log Likelihood** | -2,388.242 | -2,393.590 | -2,196.188 | -2,382.680 | -2,395.895 | -2,393.519 | -2,379.415 | -2,177.448 | -2,285.103 | -1,987.289 | -2,041.668 | -2,190.521 |
| **$sigma^2$** | 0.196 | 0.197 | 0.179 | 0.196 | 0.197 | 0.197 | 0.196 | 0.180 | 0.191 | 0.159 | 0.167 | 0.173 |
| **Akaike Inf. Crit.** | 4,784.485 | 4,795.181 | 4,400.376 | 4,773.360 | 4,799.790 | 4,795.037 | 4,766.830 | 4,362.897 | 4,578.207 | 3,982.579 | 4,091.335 | 4,389.042 |
| **Wald Test (df = 1)** | 4,276.327*** | 4,226.110*** | 3,280.112*** | 4,148.932*** | 4,223.210*** | 4,096.024*** | 4,131.984*** | 2,369.893*** | 2,473.251*** | 3,294.853*** | 2,387.532*** | 5,106.145*** |
| **LR Test (df = 1)** | 2,638.595*** | 2,627.414*** | 2,165.282*** | 2,551.976*** | 2,622.122*** | 2,472.164*** | 2,533.388*** | 1,272.201*** | 1,222.010*** | 2,126.298*** | 1,279.963*** | 2,941.016*** |

*Note:* *p **p ***p<0.001

**Bivariate Regressions for Licensed Premises Expenditures - 2**

| | *Dependent variable:* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Licensed Premises Expenditure** | | | | | | | | | | | |
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** | **(10)** | **(11)** | **(12)** |
| **Socio-Economic Variables** | | | | | | | | | | | | |
| Percentage of Apartments - Less than 5 floors | -0.002*** | | | | | | | | | | | |
| Percentage of Semi-Detached Houses | | 0.001* | | | | | | | | | | |
| Percentage of Single Detached Houses | | | 0.007*** | | | | | | | | | |
| Percentage of Row Houses | | | | 0.002*** | | | | | | | | |
| Percentage of Rented Dwellings | | | | | -0.007*** | | | | | | | |
| Percentage of Married Couples | | | | | | 0.007*** | | | | | | |
| Percentage of Lone-Parent Households | | | | | | | -0.014*** | | | | | |
| Percentage Post-Secondary Education | | | | | | | | 1.340*** | | | | |
| Average Number of Persons in Private Households | | | | | | | | | 0.157*** | | | |
| Percentage Aboriginal Identity | | | | | | | | | | -0.001 | | |
| **Built-Environment Variables** | | | | | | | | | | | | |
| Percentage of Residential Land | | | | | | | | | | | 0.104*** | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lagged Percentage of Residential Land | | | | | | | | | | | $0.236^{***}$ | |
| Percentage of Commercial Land | | | | | | | | | | | | $-0.379^{***}$ |
| Lagged Percentage of Commercial Land | | | | | | | | | | | | $-0.737^{**}$ |
| Constant | $5.571^{***}$ | $5.524^{***}$ | $5.256^{***}$ | $5.528^{***}$ | $5.781^{***}$ | $5.124^{***}$ | $5.744^{***}$ | $4.897^{***}$ | $5.108^{***}$ | $5.539^{***}$ | $5.292^{***}$ | $5.559^{***}$ |
| **Observations** | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 |
| **Log Likelihood** | -2,389.890 | -2,396.366 | -2,160.035 | -2,389.547 | -2,118.848 | -2,319.119 | -2,302.805 | -2,258.003 | -2,367.944 | -2,398.428 | -2,389.072 | -2,386.813 |
| **sigma$^2$** | 0.196 | 0.197 | 0.171 | 0.196 | 0.166 | 0.189 | 0.188 | 0.190 | 0.192 | 0.197 | 0.197 | 0.196 |
| **Akaike Inf. Crit.** | 4,787.780 | 4,800.732 | 4,328.069 | 4,787.095 | 4,245.697 | 4,646.239 | 4,613.609 | 4,524.005 | 4,743.889 | 4,804.857 | 4,788.144 | 4,783.626 |
| **Wald Test (df = 1)** | $4,255.330^{***}$ | $4,265.834^{***}$ | $4,897.484^{***}$ | $4,307.418^{***}$ | $5,166.493^{***}$ | $4,046.366^{***}$ | $3,744.761^{***}$ | $2,191.785^{***}$ | $4,889.109^{***}$ | $4,203.751^{***}$ | $4,002.365^{***}$ | $4,122.916^{***}$ |
| **LR Test (df = 1)** | $2,635.295^{***}$ | $2,606.179^{***}$ | $2,905.311^{***}$ | $2,631.549^{***}$ | $3,025.130^{***}$ | $2,547.737^{***}$ | $2,243.295^{***}$ | $1,049.947^{***}$ | $2,564.613^{***}$ | $2,617.420^{***}$ | $2,456.192^{***}$ | $2,572.257^{***}$ |

*Note:* $^{*}$p $^{**}$p $^{***}$p<0.001

**Bivariate Regressions for Licensed Premises Expenditures - 3**

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | **Licensed Premises Expenditure** | | | | | |
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| **Built-Environment Variables** | | | | | | |
| Subway Intercepts | 0.038 | | | | | |
| Lagged Subway Intercepts | 0.249** | | | | | |
| Highway Intercepts | | -0.066 | | | | |
| Lagged Highway Intercepts | | -0.166 | | | | |
| Primary Drinking Restaurant Density by Population | | | -0.752 | | | |
| Lagged Primary Drinking Restaurant Density by Population | | | 103.668** | | | |
| Primary Drinking Restaurant Density by Area (KM - Squared) | | | | -0.0003 | | |
| Lagged Primary Drinking Restaurant Density by Area (KM - Squared) | | | | 0.010* | | |
| Restaurant Density by Population | | | | | 0.063 | |
| Lagged Restaurant Density by Population | | | | | -0.001 | |
| Restaurant Density (KM - Squared) | | | | | | -0.0003 |
| Lagged Restaurant Density (KM - Squared) | | | | | | 0.001* |
| Constant | 5.521*** | 5.549*** | 5.523*** | 5.528*** | 5.539*** | 5.524*** |
| Observations | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 |
| Log Likelihood | -2,394.860 | -2,396.167 | -2,392.522 | -2,394.002 | -2,398.178 | -2,392.467 |
| $sigma^2$ | 0.198 | 0.197 | 0.197 | 0.197 | 0.197 | 0.197 |
| Akaike Inf. Crit. | 4,799.720 | 4,802.333 | 4,795.044 | 4,798.003 | 4,806.355 | 4,794.933 |
| Wald Test (df = 1) | 4,042.191*** | 4,197.819*** | 4,156.006*** | 4,148.950*** | 4,205.267*** | 4,118.310*** |
| LR Test (df = 1) | 2,454.602*** | 2,610.402*** | 2,586.508*** | 2,573.922*** | 2,618.637*** | 2,511.452*** |
| *Note:* | | | | | | *p **p ***p<0.001 |

# A3. Bivariate Spatial Error Regressions for Purchased in Stores Expenditures

**Bivariate Regressions for Purchased in Stores Expenditures - 1**

| | ***Dependent variable:*** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Purchased in Store expenditure** | | | | | | | | | | | |
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** | **(10)** | **(11)** | **(12)** |
| **Socio-Economic Variables** | | | | | | | | | | | | |
| Percentage of 20 to 44 Year Olds | -0.956*** | | | | | | | | | | | |
| Male to Female Ratio | | 0.034** | | | | | | | | | | |
| Percentage White | | | 0.011*** | | | | | | | | | |
| Percentage of Black Ethnicity | | | | -0.038*** | | | | | | | | |
| Percentage of Chinese Ethnicity | | | | | -0.002* | | | | | | | |
| Percentage of Filipino Ethnicity | | | | | | -0.003* | | | | | | |
| Percentage of South Asian Ethnicity | | | | | | | -0.027*** | | | | | |
| Percentage of Persons in White-collar Occupations | | | | | | | | 1.483*** | | | | |
| Percentage of Persons in Blue-collar Occupations | | | | | | | | | -1.387*** | | | |
| Median Household Income - After-Tax | | | | | | | | | | 0.00001*** | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Percentage of Income from Transfer Payments | | | | | | | | | | | -0.038$^{***}$ | |
| Percentage of Apartments - 5 floors or more | | | | | | | | | | | | -0.007$^{***}$ |
| Constant | 6.583$^{***}$ | 6.190$^{***}$ | 5.602$^{***}$ | 6.235$^{***}$ | 6.238$^{***}$ | 6.232$^{***}$ | 6.232$^{***}$ | 5.451$^{***}$ | 6.464$^{***}$ | 5.493$^{***}$ | 6.657$^{***}$ | 6.337$^{***}$ |
| Observations | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 |
| Log Likelihood | -1,869.874 | -1,912.404 | -1,601.255 | -1,886.237 | -1,915.503 | -1,914.398 | -1,896.828 | -1,637.364 | -1,799.629 | -1,125.912 | -1,344.475 | -1,499.459 |
| $sigma^2$ | 0.146 | 0.150 | 0.128 | 0.148 | 0.150 | 0.150 | 0.148 | 0.133 | 0.145 | 0.098 | 0.113 | 0.116 |
| Akaike Inf. Crit. | 3,747.749 | 3,832.809 | 3,210.510 | 3,780.473 | 3,839.006 | 3,836.797 | 3,801.657 | 3,282.729 | 3,607.258 | 2,259.824 | 2,696.950 | 3,006.919 |
| Wald Test (df = 1) | 4,238.056$^{***}$ | 4,281.194$^{***}$ | 3,194.526$^{***}$ | 4,192.024$^{***}$ | 4,224.912$^{***}$ | 4,158.043$^{***}$ | 4,187.849$^{***}$ | 2,174.317$^{***}$ | 2,620.805$^{***}$ | 2,655.114$^{***}$ | 2,107.889$^{***}$ | 5,432.539$^{***}$ |
| LR Test (df = 1) | 2,568.492$^{***}$ | 2,569.793$^{***}$ | 2,084.813$^{***}$ | 2,484.936$^{***}$ | 2,541.843$^{***}$ | 2,436.839$^{***}$ | 2,489.504$^{***}$ | 1,150.269$^{***}$ | 1,360.467$^{***}$ | 1,734.216$^{***}$ | 1,115.471$^{***}$ | 3,046.623$^{***}$ |
| *Note:* | | | | | | | | | | | | $^{*}$p $^{**}$p $^{***}$p<0.001 |

**Bivariate Regressions for Purchased in Stores Expenditures - 2**

| | *Dependent variable:* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Purchased in Store Expenditure** | | | | | | | | | | | |
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** | **(10)** | **(11)** | **(12)** |
| **<u>Socio-Economic Variables</u>** | | | | | | | | | | | | |
| Percentage of Apartments - Less than 5 floors | -0.002$^{***}$ | | | | | | | | | | | |
| Percentage of Semi-Detached Houses | | 0.004$^{***}$ | | | | | | | | | | |
| Percentage of Single Detached Houses | | | 0.008$^{***}$ | | | | | | | | | |
| Percentage of Row Houses | | | | 0.001 | | | | | | | | |
| Percentage of Rented Dwellings | | | | | -0.009$^{***}$ | | | | | | | |
| Percentage of Married Couples | | | | | | 0.009$^{***}$ | | | | | | |
| Percentage of Lone-Parent Households | | | | | | | -0.017$^{***}$ | | | | | |
| Percentage Post-Secondary Education | | | | | | | | 1.226$^{***}$ | | | | |
| Average Number of Persons in Private Households | | | | | | | | | 0.232$^{***}$ | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Percentage Aboriginal Identity | | | | | | | | | | -0.005 | | |
| **Built-Environment Variables** | | | | | | | | | | | | |
| Percentage of Residential Land | | | | | | | | | | | 0.132*** | |
| Lagged Percentage of Residential Land | | | | | | | | | | | 0.305*** | |
| Percentage of Commercial Land | | | | | | | | | | | | -0.411*** |
| Lagged Percentage of Commercial Land | | | | | | | | | | | | -0.846*** |
| Constant | 6.262*** | 6.175*** | 5.882*** | 6.217*** | 6.527*** | 5.672*** | 6.475*** | 5.636*** | 5.582*** | 6.223*** | 5.905*** | 6.243*** |
| Observations | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 |
| Log Likelihood | -1,899.080 | -1,890.678 | -1,438.331 | -1,916.649 | -1,245.212 | -1,728.965 | -1,722.128 | -1,753.422 | -1,827.026 | -1,916.826 | -1,897.412 | -1,899.215 |
| $sigma^2$ | 0.148 | 0.147 | 0.113 | 0.150 | 0.100 | 0.136 | 0.136 | 0.142 | 0.141 | 0.150 | 0.149 | 0.149 |
| Akaike Inf. Crit. | 3,806.160 | 3,789.355 | 2,884.662 | 3,841.299 | 2,498.424 | 3,465.929 | 3,452.257 | 3,514.844 | 3,662.052 | 3,841.652 | 3,804.823 | 3,808.429 |
| Wald Test (df = 1) | 4,270.161*** | 4,555.508*** | 4,977.815*** | 4,286.554*** | 5,804.283*** | 3,867.741*** | 3,582.776*** | 2,349.476*** | 5,279.760*** | 4,243.191*** | 3,943.123*** | 4,133.657*** |
| LR Test (df = 1) | 2,582.991*** | 2,611.834*** | 2,929.563*** | 2,537.574*** | 3,295.202*** | 2,385.696*** | 2,076.627*** | 1,243.561*** | 2,732.078*** | 2,551.926*** | 2,381.355*** | 2,502.393*** |
| *Note:* | | | | | | | | | | | | *p **p ***p<0.001 |

**Bivariate Regressions for Purchased in Stores Expenditures - 3**

| | | | | |
|---|---|---|---|---|
| | *Dependent variable:* | | | |
| | Purchased in Store Expenditure | | | |
| | (1) | (2) | (3) | (4) |

| | | | | |
|---|---|---|---|---|
| **Built-Environment Variables** | | | | |
| Subway Intercepts | -0.001 | | | |
| Lagged Subway Intercepts | 0.108 | | | |
| Highway Intercepts | | -0.050 | | |
| Lagged Highway Intercepts | | -0.202$^{*}$ | | |
| Outlet Density by Population | | | 0.009 | |
| Outlet Density by Area (KM - Squared) | | | | -0.001 |
| Lagged Outlet Density by Area (KM - Squared) | | | | 0.006 |
| Constant | 6.213$^{***}$ | 6.233$^{***}$ | 6.221$^{***}$ | 6.218$^{***}$ |
| Observations | 3,512 | 3,512 | 3,512 | 3,512 |
| Log Likelihood | -1,916.661 | -1,914.419 | -1,917.555 | -1,915.995 |
| $sigma^2$ | 0.150 | 0.150 | 0.150 | 0.150 |
| Akaike Inf. Crit. | 3,843.323 | 3,838.837 | 3,843.110 | 3,841.989 |
| Wald Test (df = 1) | 4,195.934$^{***}$ | 4,247.703$^{***}$ | 4,250.121$^{***}$ | 4,241.335$^{***}$ |
| LR Test (df = 1) | 2,470.304$^{***}$ | 2,556.123$^{***}$ | 2,559.466$^{***}$ | 2,535.984$^{***}$ |
| *Note:* | | | | $^{*}$p $^{**}$p $^{***}$p<0.001 |

## A4. Correlation Matrix

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Expend. in Licensed Premises | 1.00 | | | | | | | | | | | |
| Expend. From Purchases in Store | 0.94 | 1.00 | | | | | | | | | | |
| Log Transformed - LP | 0.91 | 0.83 | 1.00 | | | | | | | | | |
| Log Transformed - PFS | 0.86 | 0.89 | 0.93 | 1.00 | | | | | | | | |
| % 20 to 44 years | -0.03 | -0.19 | -0.01 | -0.16 | 1.00 | | | | | | | |
| Male to Female Ratio | 0.00 | -0.01 | 0.01 | 0.00 | 0.14 | 1.00 | | | | | | |
| % White | 0.41 | 0.43 | 0.47 | 0.52 | -0.11 | -0.02 | 1.00 | | | | | |
| % Ethnicity Black | -0.12 | -0.13 | -0.17 | -0.20 | 0.03 | -0.02 | -0.23 | 1.00 | | | | |
| % Ethnicity Chinese CHN | -0.07 | -0.11 | -0.02 | -0.08 | 0.07 | 0.00 | -0.50 | -0.04 | 1.00 | | | |
| % Ethnicity Filipino FLP | -0.17 | -0.16 | -0.21 | -0.19 | 0.01 | -0.02 | -0.26 | 0.08 | -0.07 | 1.00 | | |
| % Ethnicity South Asian SA | -0.14 | -0.13 | -0.19 | -0.18 | 0.02 | 0.00 | -0.29 | 0.08 | -0.02 | 0.06 | 1.00 | |
| % Blue-collar | -0.52 | -0.50 | -0.61 | -0.58 | -0.03 | 0.03 | -0.44 | 0.15 | -0.04 | 0.16 | 0.21 | 1.00 |
| % White-collar | 0.55 | 0.57 | 0.63 | 0.66 | -0.11 | -0.04 | 0.41 | -0.17 | 0.02 | -0.18 | -0.16 | -0.77 |
| Median Household Income - BT | 0.58 | 0.71 | 0.57 | 0.71 | -0.35 | 0.00 | 0.33 | -0.13 | -0.03 | -0.13 | -0.09 | -0.36 |
| Median Household Income - AT | 0.56 | 0.70 | 0.56 | 0.70 | -0.37 | 0.00 | 0.31 | -0.14 | -0.01 | -0.12 | -0.08 | -0.34 |
| % Transfer Payments | -0.56 | -0.57 | -0.67 | -0.72 | -0.07 | -0.05 | -0.41 | 0.23 | 0.01 | 0.10 | 0.17 | 0.53 |
| % Apartments – 5 or more floors | -0.08 | -0.18 | -0.16 | -0.31 | 0.25 | -0.04 | -0.22 | 0.09 | -0.03 | 0.03 | 0.08 | -0.01 |
| % Apartments – Less than 5 floors | 0.00 | -0.07 | 0.01 | -0.06 | 0.42 | 0.05 | 0.15 | 0.00 | -0.04 | -0.03 | -0.08 | -0.04 |
| % Apartments - Duplex | -0.15 | -0.10 | -0.12 | -0.02 | -0.12 | 0.01 | -0.09 | -0.01 | -0.05 | 0.13 | 0.09 | 0.23 |
| % Row Houses | -0.06 | -0.09 | -0.04 | -0.08 | 0.04 | -0.02 | -0.24 | 0.11 | 0.15 | 0.04 | 0.03 | 0.06 |
| % Semi-Detached | -0.09 | -0.07 | -0.07 | -0.02 | 0.04 | 0.00 | 0.01 | -0.05 | 0.04 | -0.09 | -0.01 | 0.14 |
| % Single Detached | 0.18 | 0.32 | 0.23 | 0.39 | -0.55 | 0.02 | 0.22 | -0.10 | 0.00 | -0.01 | -0.05 | -0.10 |
| % Dwellings Rented | -0.12 | -0.26 | -0.21 | -0.40 | 0.47 | -0.01 | -0.14 | 0.15 | -0.15 | 0.06 | 0.04 | 0.05 |



Correlation Coefficient > 0.5

Correlation Coefficient < -0.5

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % Married Couples | 0.22 | 0.32 | 0.25 | 0.38 | -0.39 | -0.03 | 0.10 | -0.21 | 0.16 | -0.04 | 0.01 | -0.11 |

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % Lone-Parent | -0.33 | -0.37 | -0.39 | -0.47 | 0.12 | 0.01 | -0.27 | 0.25 | -0.08 | 0.09 | 0.05 | 0.26 |
| % Post-Secondary Education | 0.57 | 0.53 | 0.64 | 0.61 | 0.16 | -0.01 | 0.33 | -0.15 | 0.03 | -0.10 | -0.13 | -0.72 |
| Avg. Number of Person in Private Dwelling | -0.15 | -0.02 | -0.18 | -0.05 | -0.34 | 0.07 | -0.48 | 0.08 | 0.26 | 0.10 | 0.20 | 0.37 |
| % Aboriginal Identity | -0.03 | -0.06 | -0.02 | -0.05 | 0.09 | 0.00 | 0.05 | 0.02 | -0.07 | -0.01 | -0.02 | 0.03 |
| % Residential Land | 0.09 | 0.10 | 0.13 | 0.16 | -0.02 | -0.03 | 0.10 | -0.02 | 0.05 | -0.01 | -0.04 | -0.10 |
| Lagged % Residential Land | 0.17 | 0.18 | 0.22 | 0.23 | -0.01 | -0.03 | 0.18 | -0.06 | 0.05 | -0.04 | -0.06 | -0.21 |
| % Commercial Land | -0.07 | -0.07 | -0.09 | -0.11 | 0.01 | -0.01 | -0.09 | 0.01 | 0.02 | 0.03 | 0.01 | 0.04 |
| Lagged % Commercial Land | -0.10 | -0.11 | -0.12 | -0.14 | 0.03 | 0.00 | -0.15 | 0.04 | 0.03 | 0.03 | 0.08 | 0.09 |
| Highway Intercept | -0.03 | -0.02 | -0.04 | -0.04 | -0.01 | 0.06 | -0.04 | -0.01 | -0.03 | 0.01 | 0.02 | -0.01 |
| Lagged Highway Intercept | -0.05 | -0.04 | -0.06 | -0.05 | -0.06 | 0.04 | -0.06 | -0.02 | -0.03 | 0.02 | 0.02 | 0.02 |
| Subway Intercept | 0.12 | 0.08 | 0.14 | 0.10 | 0.15 | -0.01 | 0.08 | -0.04 | -0.03 | -0.06 | -0.04 | -0.17 |
| Lagged Subway Intercept | 0.21 | 0.15 | 0.22 | 0.16 | 0.23 | -0.01 | 0.14 | -0.07 | -0.06 | -0.06 | -0.04 | -0.26 |
| Outlet Density by Population | 0.01 | 0.00 | 0.02 | 0.01 | 0.04 | 0.06 | -0.01 | 0.02 | -0.01 | -0.01 | -0.01 | 0.00 |
| Outlet Density by Area | 0.04 | 0.01 | 0.05 | 0.02 | 0.13 | 0.02 | 0.01 | 0.04 | 0.03 | -0.01 | -0.02 | -0.07 |
| Lagged Outlet Density by Area | 0.11 | 0.06 | 0.14 | 0.09 | 0.20 | 0.03 | 0.06 | -0.01 | 0.01 | -0.06 | -0.03 | -0.18 |
| Primary Drinking Restaurant Density By Area | 0.03 | -0.01 | 0.04 | 0.00 | 0.23 | 0.02 | 0.03 | -0.01 | 0.03 | -0.04 | -0.03 | -0.09 |
| Lagged Primary Drinking Restaurant Density By Area | 0.09 | 0.00 | 0.12 | 0.04 | 0.40 | 0.05 | 0.07 | 0.00 | 0.05 | -0.08 | -0.04 | -0.16 |



Correlation Coefficient > 0.5

Correlation Coefficient < -0.5

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Eating Restaurant Density by Area | 0.07 | -0.01 | 0.10 | 0.01 | 0.38 | 0.02 | 0.00 | 0.01 | 0.08 | -0.05 | -0.04 | -0.16 |
| Lagged Eating Restaurant Density by Area | 0.13 | 0.02 | 0.18 | 0.07 | 0.49 | 0.03 | 0.04 | -0.01 | 0.08 | -0.09 | -0.06 | -0.23 |
| Pop Density by Area KMS | -0.06 | -0.13 | -0.08 | -0.19 | 0.34 | -0.01 | -0.17 | 0.08 | 0.03 | 0.03 | 0.05 | -0.01 |
| Lagged Pop Density by Area KMS | 0.00 | -0.11 | 0.02 | -0.11 | 0.45 | 0.02 | -0.09 | 0.04 | 0.04 | 0.02 | 0.00 | -0.09 |
| Distance to Nearest Outlet | -0.13 | -0.07 | -0.15 | -0.10 | -0.24 | 0.00 | -0.16 | 0.04 | 0.02 | 0.09 | 0.07 | 0.22 |

| | *13* | *14* | *15* | *16* | *17* | *18* | *19* | *20* | *21* | *22* | *23* | *24* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % White-collar | 1.00 | | | | | | | | | | | |
| Median Household Income - BT | 0.51 | 1.00 | | | | | | | | | | |
| Median Household Income - AT | 0.50 | 0.97 | 1.00 | | | | | | | | | |
| % Transfer Payments | -0.61 | -0.51 | -0.51 | 1.00 | | | | | | | | |
| % Apartments – 5 or more floors | -0.03 | -0.35 | -0.39 | 0.21 | 1.00 | | | | | | | |
| % Apartments – Less than 5 floors | -0.16 | -0.24 | -0.27 | -0.01 | -0.27 | 1.00 | | | | | | |
| % Apartments - Duplex | -0.16 | 0.02 | 0.05 | 0.03 | -0.38 | -0.14 | 1.00 | | | | | |
| % Row Houses | -0.10 | -0.07 | -0.07 | 0.10 | -0.08 | -0.04 | -0.15 | 1.00 | | | | |
| % Semi-Detached | -0.13 | -0.02 | 0.00 | 0.04 | -0.26 | 0.12 | -0.08 | -0.02 | 1.00 | | | |
| % Single Detached | 0.28 | 0.55 | 0.58 | -0.26 | -0.55 | -0.44 | 0.34 | -0.25 | -0.28 | 1.00 | | |
| % Dwellings Rented | -0.23 | -0.54 | -0.58 | 0.29 | 0.63 | 0.35 | -0.33 | 0.03 | -0.19 | -0.70 | 1.00 | |
| % Married Couples | 0.30 | 0.46 | 0.49 | -0.28 | -0.27 | -0.31 | 0.15 | -0.14 | 0.01 | 0.50 | -0.54 | 1.00 |
| % Lone-Parent | -0.37 | -0.40 | -0.42 | 0.43 | 0.21 | 0.13 | -0.09 | 0.17 | -0.01 | -0.34 | 0.40 | -0.86 |
| % Post-Secondary Education | 0.72 | 0.42 | 0.40 | -0.65 | 0.10 | -0.01 | -0.22 | -0.06 | -0.13 | 0.05 | -0.02 | 0.18 |
| Avg. Number of Person in Private Dwelling | -0.21 | 0.28 | 0.34 | 0.14 | -0.42 | -0.31 | 0.32 | 0.16 | 0.13 | 0.43 | -0.50 | 0.34 |
| % Aboriginal Identity | -0.08 | -0.11 | -0.12 | 0.02 | 0.00 | 0.14 | -0.03 | -0.01 | 0.03 | -0.10 | 0.13 | -0.21 |
| % Residential Land | 0.09 | 0.14 | 0.15 | -0.11 | -0.24 | 0.02 | 0.12 | -0.08 | 0.07 | 0.19 | -0.17 | 0.10 |



Correlation Coefficient > 0.5

Correlation Coefficient < -0.5

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lagged % Residential Land | 0.17 | 0.16 | 0.16 | -0.19 | -0.20 | 0.06 | 0.05 | -0.07 | 0.03 | 0.15 | -0.10 | 0.07 |

| | *13* | *14* | *15* | *16* | *17* | *18* | *19* | *20* | *21* | *22* | *23* | *24* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % Commercial Land | -0.04 | -0.09 | -0.10 | 0.08 | 0.17 | -0.03 | -0.06 | 0.04 | -0.06 | -0.12 | 0.11 | -0.05 |
| Lagged % Commercial Land | -0.06 | -0.10 | -0.11 | 0.11 | 0.16 | -0.06 | -0.03 | 0.06 | -0.05 | -0.11 | 0.11 | -0.03 |
| Highway Intercept | 0.01 | -0.02 | -0.01 | 0.02 | 0.08 | -0.06 | -0.03 | 0.02 | -0.04 | -0.03 | 0.02 | -0.02 |
| Lagged Highway Intercept | 0.01 | 0.00 | 0.01 | 0.03 | 0.08 | -0.10 | -0.03 | 0.02 | -0.04 | 0.01 | -0.02 | 0.05 |
| Subway Intercept | 0.13 | -0.02 | -0.03 | -0.11 | 0.13 | 0.07 | -0.08 | -0.02 | -0.05 | -0.13 | 0.12 | -0.02 |
| Lagged Subway Intercept | 0.21 | 0.01 | -0.01 | -0.16 | 0.18 | 0.09 | -0.12 | -0.06 | -0.04 | -0.17 | 0.16 | -0.03 |
| Outlet Density by Population | -0.01 | -0.01 | -0.02 | -0.03 | 0.02 | 0.00 | 0.00 | -0.01 | 0.00 | -0.02 | 0.04 | -0.07 |
| Outlet Density by Area | 0.05 | -0.04 | -0.04 | -0.03 | 0.08 | 0.03 | -0.06 | -0.01 | -0.01 | -0.08 | 0.05 | -0.05 |
| Lagged Outlet Density by Area | 0.11 | -0.02 | -0.04 | -0.10 | 0.13 | 0.06 | -0.12 | 0.03 | -0.03 | -0.13 | 0.10 | -0.08 |
| Primary Drinking Restaurant Density By Area | 0.02 | -0.10 | -0.11 | 0.00 | 0.09 | 0.13 | -0.08 | 0.02 | -0.03 | -0.15 | 0.15 | -0.11 |
| Lagged Primary Drinking Restaurant Density By Area | 0.03 | -0.12 | -0.14 | -0.05 | 0.14 | 0.22 | -0.14 | 0.03 | -0.03 | -0.26 | 0.22 | -0.20 |
| Eating Restaurant Density by Area | 0.06 | -0.16 | -0.18 | -0.04 | 0.23 | 0.17 | -0.15 | -0.02 | -0.06 | -0.28 | 0.25 | -0.17 |
| Lagged Eating Restaurant Density by Area | 0.11 | -0.15 | -0.18 | -0.10 | 0.27 | 0.20 | -0.18 | 0.00 | -0.04 | -0.34 | 0.29 | -0.21 |
| Pop Density by Area KMS | -0.07 | -0.23 | -0.25 | 0.13 | 0.46 | 0.02 | -0.20 | -0.02 | -0.07 | -0.38 | 0.38 | -0.21 |
| Lagged Pop Density by Area KMS | -0.04 | -0.27 | -0.29 | 0.04 | 0.37 | 0.17 | -0.18 | 0.02 | -0.01 | -0.45 | 0.39 | -0.27 |
| Distance to Nearest Outlet | -0.09 | 0.05 | 0.07 | 0.09 | -0.09 | -0.22 | 0.21 | -0.06 | -0.03 | 0.23 | -0.17 | 0.14 |



Correlation Coefficient > 0.5

Correlation Coefficient < -0.5

| | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % Lone-Parent | 1.00 | | | | | | | | | | | |
| % Post-Secondary Education | -0.35 | 1.00 | | | | | | | | | | |
| Avg. Number of Person in Private Dwelling | -0.05 | -0.31 | 1.00 | | | | | | | | | |
| % Aboriginal Identity | 0.15 | -0.03 | -0.15 | 1.00 | | | | | | | | |
| % Residential Land | -0.11 | 0.07 | 0.07 | -0.03 | 1.00 | | | | | | | |
| Lagged % Residential Land | -0.11 | 0.16 | -0.03 | 0.03 | 0.37 | 1.00 | | | | | | |
| % Commercial Land | 0.05 | -0.03 | -0.06 | -0.02 | -0.27 | -0.07 | 1.00 | | | | | |
| Lagged % Commercial Land | 0.05 | -0.04 | 0.00 | -0.03 | -0.08 | -0.24 | 0.15 | 1.00 | | | | |
| Highway Intercept | 0.04 | 0.01 | 0.00 | -0.04 | -0.21 | -0.16 | 0.01 | 0.00 | 1.00 | | | |
| Lagged Highway Intercept | 0.00 | -0.01 | 0.06 | -0.04 | -0.13 | -0.24 | 0.01 | -0.01 | 0.60 | 1.00 | | |
| Subway Intercept | -0.06 | 0.19 | -0.21 | 0.01 | -0.05 | 0.02 | 0.05 | 0.03 | 0.09 | 0.02 | 1.00 | |
| Lagged Subway Intercept | -0.10 | 0.30 | -0.29 | 0.00 | -0.02 | 0.03 | 0.04 | 0.08 | 0.04 | 0.04 | 0.55 | 1.00 |
| Outlet Density by Population | 0.06 | 0.05 | -0.04 | 0.13 | -0.07 | 0.00 | 0.03 | 0.00 | 0.02 | 0.01 | 0.03 | 0.03 |
| Outlet Density by Area | 0.01 | 0.07 | -0.10 | 0.01 | -0.02 | 0.00 | 0.06 | 0.04 | 0.00 | 0.00 | 0.05 | 0.10 |
| Lagged Outlet Density by Area | -0.04 | 0.17 | -0.18 | 0.03 | -0.01 | 0.00 | 0.05 | 0.08 | 0.00 | -0.02 | 0.16 | 0.20 |
| Primary Drinking Restaurant Density By Area | 0.01 | 0.07 | -0.18 | 0.05 | -0.01 | 0.01 | -0.02 | -0.01 | -0.03 | -0.05 | 0.06 | 0.12 |
| Lagged Primary Drinking Restaurant Density By Area | 0.01 | 0.13 | -0.30 | 0.08 | -0.01 | 0.02 | 0.00 | -0.03 | -0.05 | -0.10 | 0.15 | 0.18 |
| Eating Restaurant Density by Area | 0.02 | 0.15 | -0.32 | 0.03 | -0.03 | -0.01 | 0.14 | 0.06 | -0.05 | -0.08 | 0.24 | 0.28 |
| Lagged Eating Restaurant Density by Area | 0.01 | 0.23 | -0.40 | 0.07 | -0.02 | 0.01 | 0.04 | 0.10 | -0.05 | -0.10 | 0.25 | 0.36 |
| Pop Density by Area KMS | 0.14 | 0.06 | -0.21 | 0.04 | 0.04 | -0.03 | 0.04 | 0.08 | -0.06 | -0.04 | 0.04 | 0.14 |
| Lagged Pop Density by Area KMS | 0.11 | 0.13 | -0.33 | 0.10 | -0.02 | 0.07 | 0.07 | 0.08 | -0.05 | -0.10 | 0.14 | 0.22 |



Correlation Coefficient > 0.5

Correlation Coefficient < -0.5

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distance to Nearest Outlet | 0.00 | -0.18 | 0.28 | -0.09 | 0.03 | 0.02 | -0.02 | 0.00 | 0.05 | 0.08 | -0.11 | -0.20 |

| | *37* | *38* | *39* | *40* | *41* | *42* | *43* | *44* | *45* | *46* |
|---|---|---|---|---|---|---|---|---|---|---|
| Outlet Density by Population | 1.00 | | | | | | | | | |
| Outlet Density by Area | 0.28 | 1.00 | | | | | | | | |
| Lagged Outlet Density by Area | 0.01 | 0.13 | 1.00 | | | | | | | |
| Primary Drinking Restaurant Density By Area | 0.03 | 0.10 | 0.15 | 1.00 | | | | | | |
| Lagged Primary Drinking Restaurant Density By Area | 0.02 | 0.13 | 0.27 | 0.27 | 1.00 | | | | | |
| Eating Restaurant Density by Area | 0.11 | 0.28 | 0.34 | 0.38 | 0.46 | 1.00 | | | | |
| Lagged Eating Restaurant Density by Area | 0.04 | 0.19 | 0.47 | 0.33 | 0.66 | 0.66 | 1.00 | | | |
| Pop Density by Area KMS | -0.03 | 0.07 | 0.11 | 0.08 | 0.17 | 0.21 | 0.26 | 1.00 | | |
| Lagged Pop Density by Area KMS | 0.00 | 0.09 | 0.24 | 0.17 | 0.32 | 0.32 | 0.46 | 0.40 | 1.00 | |
| Distance to Nearest Outlet | -0.08 | -0.20 | -0.34 | -0.13 | -0.23 | -0.23 | -0.30 | -0.14 | -0.24 | 1.00 |



Correlation Coefficient > 0.5

Correlation Coefficient < -0.5

# A5. Multivariate OLS Regression Models for Licensed Premises Expenditures

**Licensed Premises Expenditure - OLS Regressions**

| | ***Dependent variable:*** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Licensed Premises Expenditure** | | | | | | | | | | | |
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** | **(10)** | **(11)** | **(12)** |
| **1. Percentage of 0 to 19 Year Olds** | -1.044$^{***}$ | -1.067$^{***}$ | | | -1.022$^{***}$ | -1.045$^{***}$ | -1.024$^{***}$ | -1.047$^{***}$ | -0.975$^{***}$ | -0.983$^{***}$ | -0.924$^{***}$ | -0.957$^{***}$ |
| | (0.172) | (0.172) | | | (0.171) | (0.172) | (0.171) | (0.172) | (0.163) | (0.163) | (0.161) | (0.163) |
| **Percentage of 20 to 24 Year Olds** | -0.666$^{**}$ | -0.661$^{**}$ | | | -0.757$^{**}$ | -0.747$^{**}$ | -0.771$^{**}$ | -0.761$^{**}$ | -0.794$^{**}$ | -0.794$^{**}$ | -0.680$^{**}$ | -0.681$^{**}$ |
| | (0.319) | (0.319) | | | (0.319) | (0.320) | (0.319) | (0.320) | (0.316) | (0.317) | (0.311) | (0.311) |
| **Percentage of 25 to 44 Year Olds** | -0.006 | 0.022 | | | -0.036 | -0.009 | -0.056 | -0.028 | | | | |
| | (0.137) | (0.138) | | | (0.137) | (0.138) | (0.138) | (0.138) | | | | |
| **Percentage of 45 to 64 Year Olds** | -0.379$^{*}$ | -0.344$^{*}$ | | | -0.393$^{**}$ | -0.359$^{*}$ | -0.392$^{**}$ | -0.357$^{*}$ | -0.368$^{**}$ | -0.351$^{*}$ | -0.332$^{*}$ | -0.361$^{**}$ |
| | (0.195) | (0.195) | | | (0.195) | (0.195) | (0.194) | (0.195) | (0.181) | (0.181) | (0.178) | (0.181) |
| **Male-Female Ratio** | 0.002 | -0.002 | | | 0.006 | 0.002 | 0.006 | 0.002 | | | | |
| | (0.013) | (0.013) | | | (0.013) | (0.013) | (0.013) | (0.013) | | | | |
| **Percentage of Black Ethnicity** | 0.008 | 0.008 | | | 0.007 | 0.007 | 0.007 | 0.006 | 0.007 | 0.007 | | |
| | (0.006) | (0.006) | | | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | | |
| **Percentage of Chinese Ethnicity** | -0.001 | -0.001 | | | -0.001 | -0.001 | -0.001 | -0.001 | | | | |
| | (0.001) | (0.001) | | | (0.001) | (0.001) | (0.001) | (0.001) | | | | |
| **Percentage of Filipino Ethnicity** | -0.010$^{***}$ | -0.011$^{***}$ | | | -0.010$^{***}$ | -0.010$^{***}$ | -0.010$^{***}$ | -0.011$^{***}$ | -0.010$^{***}$ | -0.010$^{***}$ | -0.010$^{***}$ | -0.010$^{***}$ |
| | (0.001) | (0.001) | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Percentage of South Asian Ethnicity** | -0.008 | -0.008 | | | -0.007 | -0.008 | -0.007 | -0.008 | -0.007 | -0.007 | | |
| | (0.005) | (0.005) | | | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Percentage of Persons in Blue-collar Occupations** | -0.850*** | -0.863*** | -0.827*** | -0.840*** | -0.821*** | -0.833*** | -0.807*** | -0.815*** | -0.840*** | -0.845*** |
| | (0.108) | (0.109) | (0.108) | (0.109) | (0.108) | (0.109) | (0.107) | (0.107) | (0.106) | (0.106) |
| **Percentage of Persons in White-collar Occupations** | 0.289*** | 0.302*** | 0.293*** | 0.307*** | 0.291*** | 0.305*** | 0.291*** | 0.304*** | 0.299*** | 0.286*** |
| | (0.092) | (0.092) | (0.092) | (0.092) | (0.091) | (0.092) | (0.091) | (0.091) | (0.088) | (0.090) |
| **Median Household Income - After-Tax** | 0.00001*** | | 0.00001*** | | 0.00001*** | | 0.00001*** | | 0.00001*** | |
| | (0.00000) | | (0.00000) | | (0.00000) | | (0.00000) | | (0.00000) | |
| **Median Household Income - Before-Tax** | | 0.00001*** | | 0.00001*** | | 0.00001*** | | 0.00001*** | | 0.00001*** |
| | | (0.00000) | | (0.00000) | | (0.00000) | | (0.00000) | | (0.00000) |
| **Percentage of Income from Transfer Payments** | -0.024*** | -0.024*** | -0.024*** | -0.024*** | -0.024*** | -0.024*** | -0.024*** | -0.024*** | -0.024*** | -0.024*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| **Percentage of Apartments - 5 floors or more** | 0.005 | 0.004 | 0.004 | 0.003 | 0.004 | 0.003 | -0.002*** | -0.002*** | -0.002*** | -0.002*** |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.0004) | (0.0004) | (0.0003) | (0.0003) |
| **Percentage of Apartments - Less than 5 floors** | 0.007 | 0.006 | 0.006 | 0.005 | 0.006 | 0.005 | | | | |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | | | | |
| **Percentage of Apartments - Duplexes** | 0.005 | 0.005 | 0.004 | 0.004 | 0.004 | 0.004 | -0.001 | -0.001 | | |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.001) | (0.001) | | |
| **Percentage of Row Houses** | 0.010 | 0.009 | 0.009 | 0.009 | 0.009 | 0.009 | 0.003*** | 0.003*** | 0.003*** | 0.004*** |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (0.008) | (0.008) | | (0.008) | (0.008) | (0.008) | (0.008) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Percentage of Semi-Detached Houses** | 0.006 | 0.006 | | 0.005 | 0.005 | 0.005 | 0.005 | -0.0002 | -0.00004 | | |
| | (0.008) | (0.008) | | (0.008) | (0.008) | (0.008) | (0.008) | (0.001) | (0.001) | | |
| **Percentage of Single Detached Houses** | 0.007 | 0.007 | | 0.006 | 0.006 | 0.006 | 0.006 | 0.001 | 0.001 | | 0.001* |
| | (0.008) | (0.008) | | (0.008) | (0.008) | (0.008) | (0.008) | (0.001) | (0.001) | | (0.0004) |
| **Percentage of Rented Dwellings** | 0.003*** | 0.003*** | | 0.003*** | 0.003*** | 0.003*** | 0.003*** | 0.003*** | 0.003*** | 0.003*** | 0.003*** |
| | (0.001) | (0.001) | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.0005) | (0.0004) | (0.0004) |
| **Percentage of Married Couples** | -0.001 | -0.001 | | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | | |
| | (0.001) | (0.001) | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | | |
| **Percentage of Female Parents** | -0.005*** | -0.005*** | | -0.005*** | -0.005*** | -0.005*** | -0.005*** | -0.005*** | -0.005*** | -0.004*** | -0.004*** |
| | (0.002) | (0.002) | | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) |
| **Percentage Post-Secondary Education** | 0.581*** | 0.599*** | | 0.563*** | 0.582*** | 0.566*** | 0.585*** | 0.554*** | 0.576*** | 0.542*** | 0.575*** |
| | (0.082) | (0.082) | | (0.082) | (0.082) | (0.082) | (0.082) | (0.080) | (0.080) | (0.079) | (0.079) |
| **Average Number of Persons in Private Households** | -0.067** | -0.036 | | -0.061** | -0.030 | -0.060** | -0.029 | -0.070*** | -0.043* | -0.081*** | -0.057** |
| | (0.028) | (0.028) | | (0.028) | (0.028) | (0.028) | (0.028) | (0.025) | (0.025) | (0.024) | (0.024) |
| **Percentage Aboriginal Identity** | 0.004 | 0.003 | | 0.004 | 0.003 | 0.004 | 0.003 | | | | |
| | (0.004) | (0.004) | | (0.004) | (0.004) | (0.004) | (0.004) | | | | |
| **Population Density By Area (KM - Squared)** | 0.00000** | 0.00000** | | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | | |
| | (0.00000) | (0.00000) | | (0.00000) | (0.00000) | (0.00000) | (0.00000) | (0.00000) | (0.00000) | | |
| **Outlet Density by Population** | | | 0.086 | 0.016 | 0.017 | | | | | | |
| | | | (0.061) | (0.039) | (0.039) | | | | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Outlet Density by Area (KM - Squared)** | | | | 0.008*** | | | 0.003* | 0.003* | 0.003* | 0.003* | 0.003* | 0.003* |
| | | | | (0.003) | | | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| **Percentage of Residential Land** | | | 0.261*** | 0.257*** | 0.057** | 0.058** | 0.057** | 0.057** | 0.053** | 0.054** | 0.058** | 0.058** |
| | | | (0.038) | (0.038) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) | (0.024) | (0.025) |
| **Percentage of Commercial Land** | | | -0.428*** | -0.444*** | -0.118* | -0.115 | -0.124* | -0.122* | -0.128* | -0.126* | -0.132* | -0.130* |
| | | | (0.113) | (0.113) | (0.071) | (0.071) | (0.071) | (0.071) | (0.071) | (0.071) | (0.071) | (0.071) |
| **Highway Intercepts** | | | -0.093* | -0.091* | -0.073** | -0.066* | -0.073** | -0.066* | -0.073** | -0.066* | -0.076** | -0.070** |
| | | | (0.054) | (0.054) | (0.034) | (0.034) | (0.034) | (0.034) | (0.034) | (0.034) | (0.034) | (0.034) |
| **Subway Intercepts** | | | 0.446*** | 0.440*** | 0.094*** | 0.089*** | 0.093*** | 0.089*** | 0.093*** | 0.090*** | 0.087*** | 0.085*** |
| | | | (0.049) | (0.049) | (0.032) | (0.032) | (0.032) | (0.032) | (0.032) | (0.032) | (0.032) | (0.032) |
| **Constant** | 4.874*** | 4.963*** | 5.354*** | 5.353*** | 4.944*** | 5.027*** | 4.946*** | 5.030*** | 5.515*** | 5.561*** | 5.463*** | 5.519*** |
| | (0.801) | (0.802) | (0.032) | (0.032) | (0.800) | (0.801) | (0.799) | (0.801) | (0.116) | (0.116) | (0.106) | (0.106) |
| **Observations** | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 |
| **$R^2$** | 0.631 | 0.629 | 0.044 | 0.046 | 0.633 | 0.632 | 0.634 | 0.632 | 0.633 | 0.632 | 0.632 | 0.631 |
| **Adjusted $R^2$** | 0.628 | 0.627 | 0.043 | 0.045 | 0.630 | 0.629 | 0.630 | 0.629 | 0.631 | 0.629 | 0.630 | 0.629 |
| **Residual Std. Error** | 0.424 (df = 3485) | 0.425 (df = 3485) | 0.680 (df = 3506) | 0.680 (df = 3506) | 0.423 (df = 3480) | 0.424 (df = 3480) | 0.423 (df = 3480) | 0.424 (df = 3480) | 0.423 (df = 3485) | 0.424 (df = 3485) | 0.423 (df = 3492) | 0.424 (df = 3491) |
| **F Statistic** | 229.100*** (df = 26; 3485) | 227.707*** (df = 26; 3485) | 32.424*** (df = 5; 3506) | 33.747*** (df = 5; 3506) | 193.950*** (df = 31; 3480) | 192.633*** (df = 31; 3480) | 194.228*** (df = 31; 3480) | 192.899*** (df = 31; 3480) | 231.609*** (df = 26; 3485) | 230.016*** (df = 26; 3485) | 316.009*** (df = 19; 3492) | 298.372*** (df = 20; 3491) |

*Notes:*
Significance: *p<0.1; **p<0.05; ***p<0.01
Model 1 - All socio-economic variables - After Tax Income
Model 2 - All socio-economic variables - Before - Tax Income
Model 3 - All built-environment variables - Outlet Den. by Pop
Model 4 - All built-environment variables - Outlet Den. by Area
Model 5 - All socio-economic and built-environment - After-Tax - Outlet Den by Pop
Model 6 - All socio-economic and built-environment - Before-Tax - Outlet Dent by Pop
Model 7 - All socio-economic and built-environment - After-Tax - Outlet Den by Area
Model 8 - All socio-economic and built-environment - Before -Tax - Outlet Den by Area
Model 9 - All significant variables - After - Tax Income
Model 10 - All significant variables - Before - Tax Income
Model 11 - Backwards Regression - 0.1 alpha - After – Tax
Model 12 - Backwards Regression - 0.1 alpha - Before-Tax

# A6. Multivariate OLS Regression Models for Purchased in Stores Expenditures

**Purchased From Stores Expenditures - OLS Regressions**

| | *Dependent variable:* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Purchased from Stores Expenditure | | | | | | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| 1. Percentage of 0 to 19 Year Olds | -0.810$^{***}$ | -0.834$^{***}$ | | | -0.797$^{***}$ | -0.821$^{***}$ | -0.798$^{***}$ | -0.822$^{***}$ | -0.808$^{***}$ | -0.823$^{***}$ | -0.855$^{***}$ | -0.757$^{***}$ |
| | (0.123) | (0.123) | | | (0.123) | (0.123) | (0.123) | (0.123) | (0.121) | (0.122) | (0.095) | (0.094) |
| Percentage of 20 to 24 Year Olds | -0.727$^{***}$ | -0.723$^{***}$ | | | -0.788$^{***}$ | -0.780$^{***}$ | -0.797$^{***}$ | -0.789$^{***}$ | -0.812$^{***}$ | -0.807$^{***}$ | -0.870$^{***}$ | -0.737$^{***}$ |
| | (0.228) | (0.229) | | | (0.229) | (0.229) | (0.228) | (0.229) | (0.228) | (0.228) | (0.210) | (0.209) |
| Percentage of 25 to 44 Year Olds | -0.387$^{***}$ | -0.360$^{***}$ | | | -0.405$^{***}$ | -0.378$^{***}$ | -0.420$^{***}$ | -0.393$^{***}$ | -0.430$^{***}$ | -0.408$^{***}$ | -0.421$^{***}$ | -0.418$^{***}$ |
| | (0.098) | (0.099) | | | (0.098) | (0.099) | (0.099) | (0.099) | (0.097) | (0.098) | (0.088) | (0.086) |
| Percentage of 45 to 64 Year Olds | -0.307$^{**}$ | -0.275$^{**}$ | | | -0.314$^{**}$ | -0.282$^{**}$ | -0.313$^{**}$ | -0.282$^{**}$ | -0.299$^{**}$ | -0.271$^{*}$ | -0.307$^{**}$ | -0.297$^{**}$ |
| | (0.139) | (0.140) | | | (0.139) | (0.140) | (0.139) | (0.140) | (0.139) | (0.139) | (0.138) | (0.138) |
| Male-Female Ratio | -0.007 | -0.011 | | | -0.005 | -0.009 | -0.005 | -0.009 | | | | |
| | (0.009) | (0.009) | | | (0.009) | (0.009) | (0.009) | (0.009) | | | | |
| Percentage of Black Ethnicity | 0.004 | 0.003 | | | 0.004 | 0.003 | 0.003 | 0.003 | 0.004 | 0.004 | | |
| | (0.004) | (0.004) | | | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | | |
| Percentage of Chinese Ethnicity | -0.003$^{***}$ | -0.004$^{***}$ | | | -0.003$^{***}$ | -0.004$^{***}$ | -0.003$^{***}$ | -0.004$^{***}$ | -0.003$^{***}$ | -0.004$^{***}$ | -0.003$^{***}$ | -0.003$^{***}$ |
| | (0.0004) | (0.0004) | | | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0003) |
| Percentage of Filipino Ethnicity | -0.007$^{***}$ | -0.007$^{***}$ | | | -0.007$^{***}$ | -0.007$^{***}$ | -0.007$^{***}$ | -0.007$^{***}$ | -0.006$^{***}$ | -0.007$^{***}$ | -0.006$^{***}$ | -0.007$^{***}$ |
| | (0.001) | (0.001) | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Percentage of South Asian Ethnicity | -0.005 | -0.005 | | | -0.005 | -0.005 | -0.005 | -0.005 | -0.004 | -0.005 | | |
| | (0.003) | (0.003) | | | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Percentage of Persons in Blue-collar Occupations** | -0.624$^{***}$ | -0.636$^{***}$ | | -0.610$^{***}$ | -0.622$^{***}$ | -0.605$^{***}$ | -0.616$^{***}$ | -0.573$^{***}$ | -0.586$^{***}$ | -0.584$^{***}$ | -0.568$^{***}$ |
| | (0.078) | (0.078) | | (0.078) | (0.078) | (0.078) | (0.078) | (0.078) | (0.078) | (0.075) | (0.074) |
| **Percentage of Persons in White-collar Occupations** | 0.266$^{***}$ | 0.278$^{***}$ | | 0.269$^{***}$ | 0.281$^{***}$ | 0.267$^{***}$ | 0.280$^{***}$ | 0.283$^{***}$ | 0.295$^{***}$ | 0.281$^{***}$ | 0.291$^{***}$ |
| | (0.066) | (0.066) | | (0.066) | (0.066) | (0.066) | (0.066) | (0.065) | (0.066) | (0.064) | (0.064) |
| **Median Household Income - After-Tax** | 0.00001$^{***}$ | | | 0.00001$^{***}$ | | 0.00001$^{***}$ | | 0.00001$^{***}$ | | 0.00001$^{***}$ | |
| | (0.00000) | | | (0.00000) | | (0.00000) | | (0.00000) | | (0.00000) | |
| **Median Household Income - Before-Tax** | | 0.00001$^{***}$ | | | 0.00001$^{***}$ | | 0.00001$^{***}$ | | 0.00001$^{***}$ | | 0.00001$^{***}$ |
| | | (0.00000) | | | (0.00000) | | (0.00000) | | (0.00000) | | (0.00000) |
| **Percentage of Income from Transfer Payments** | -0.024$^{***}$ | -0.024$^{***}$ | | -0.024$^{***}$ | -0.024$^{***}$ | -0.024$^{***}$ | -0.024$^{***}$ | -0.024$^{***}$ | -0.024$^{***}$ | -0.024$^{***}$ | -0.024$^{***}$ |
| | (0.001) | (0.001) | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Percentage of Apartments - 5 floors or more** | -0.001 | -0.002 | | -0.002 | -0.002 | -0.002 | -0.002 | -0.002$^{***}$ | -0.002$^{***}$ | -0.002$^{***}$ | -0.002$^{***}$ |
| | (0.006) | (0.006) | | (0.006) | (0.006) | (0.006) | (0.006) | (0.0004) | (0.0004) | (0.0002) | (0.0002) |
| **Percentage of Apartments - Less than 5 floors** | 0.001 | 0.0003 | | 0.0003 | -0.0002 | 0.0003 | -0.0002 | -0.0001 | -0.0002 | | |
| | (0.006) | (0.006) | | (0.006) | (0.006) | (0.006) | (0.006) | (0.0005) | (0.0005) | | |
| **Percentage of Apartments - Duplexes** | 0.001 | 0.001 | | 0.0004 | -0.00002 | 0.0004 | -0.00000 | | | | |
| | (0.006) | (0.006) | | (0.006) | (0.006) | (0.006) | (0.006) | | | | |
| **Percentage of Row Houses** | 0.002 | 0.001 | | 0.001 | 0.001 | 0.001 | 0.001 | 0.001$^{**}$ | 0.001$^{*}$ | 0.001$^{***}$ | 0.001$^{***}$ |
| | (0.006) | (0.006) | | (0.006) | (0.006) | (0.006) | (0.006) | (0.0005) | (0.0005) | (0.0004) | (0.0004) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Percentage of Semi-Detached Houses** | 0.001 | 0.001 | | | 0.0002 | -0.0001 | 0.0002 | -0.0001 | | | | |
| | (0.006) | (0.006) | | | (0.006) | (0.006) | (0.006) | (0.006) | | | | |
| **Percentage of Single Detached Houses** | 0.001 | 0.0002 | | | -0.0001 | -0.0004 | -0.0001 | -0.0004 | -0.0003 | -0.0003 | | |
| | (0.006) | (0.006) | | | (0.006) | (0.006) | (0.006) | (0.006) | (0.0004) | (0.0004) | | |
| **Percentage of Rented Dwellings** | 0.0004 | 0.0002 | | | 0.0003 | 0.0001 | 0.0004 | 0.0002 | 0.0004 | 0.0002 | 0.001* | |
| | (0.0004) | (0.0004) | | | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0003) | |
| **Percentage of Married Couples** | 0.001 | 0.0005 | | | 0.001 | 0.0004 | 0.0005 | 0.0004 | 0.001 | 0.0004 | | |
| | (0.001) | (0.001) | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | | |
| **Percentage of Female Parents** | -0.004*** | -0.004*** | | | -0.004*** | -0.004*** | -0.004*** | -0.004*** | -0.004*** | -0.004*** | -0.004*** | -0.005*** |
| | (0.001) | (0.001) | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Percentage Post-Secondary Education** | 0.400*** | 0.416*** | | | 0.388*** | 0.405*** | 0.390*** | 0.407*** | 0.383*** | 0.401*** | 0.387*** | 0.400*** |
| | (0.059) | (0.059) | | | (0.059) | (0.059) | (0.059) | (0.059) | (0.058) | (0.058) | (0.058) | (0.058) |
| **Average Number of Persons in Private Households** | -0.020 | 0.009 | | | -0.016 | 0.013 | -0.015 | 0.014 | -0.010 | 0.017 | | |
| | (0.020) | (0.020) | | | (0.020) | (0.020) | (0.020) | (0.020) | (0.020) | (0.019) | | |
| **Percentage Aboriginal Identity** | 0.001 | 0.001 | | | 0.001 | 0.0004 | 0.001 | 0.001 | 0.001 | 0.0001 | | |
| | (0.003) | (0.003) | | | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | | |
| **Population Density By Area (KM - Squared)** | 0.00000 | 0.00000 | | | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | | |
| | (0.00000) | (0.00000) | | | (0.00000) | (0.00000) | (0.00000) | (0.00000) | (0.00000) | (0.00000) | | |
| **Outlet Density by Population** | | | 0.049 | | 0.014 | 0.015 | | | | | | |
| | | | (0.053) | | (0.028) | (0.028) | | | | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Outlet Density by Area (KM - Squared)** | | | | 0.003 | | | 0.003$^{*}$ | 0.002$^{*}$ | | | | |
| | | | | (0.003) | | | (0.001) | (0.001) | | | | |
| **Percentage of Residential Land** | | | 0.262$^{***}$ | 0.260$^{***}$ | 0.035$^{**}$ | 0.036$^{**}$ | 0.035$^{*}$ | 0.035$^{**}$ | 0.038$^{**}$ | 0.038$^{**}$ | 0.044$^{**}$ | 0.050$^{***}$ |
| | | | (0.033) | (0.033) | (0.018) | (0.018) | (0.018) | (0.018) | (0.018) | (0.018) | (0.017) | (0.017) |
| **Percentage of Commercial Land** | | | -0.427$^{***}$ | -0.433$^{***}$ | -0.065 | -0.062 | -0.069 | -0.067 | -0.065 | -0.063 | | |
| | | | (0.098) | (0.098) | (0.051) | (0.051) | (0.051) | (0.051) | (0.051) | (0.051) | | |
| **Highway Intercepts** | | | -0.048 | -0.047 | -0.046$^{*}$ | -0.039 | -0.046$^{*}$ | -0.039 | -0.042$^{*}$ | -0.036 | -0.042$^{*}$ | |
| | | | (0.047) | (0.047) | (0.024) | (0.024) | (0.024) | (0.024) | (0.024) | (0.024) | (0.024) | |
| **Subway Intercepts** | | | 0.280$^{***}$ | 0.278$^{***}$ | 0.063$^{***}$ | 0.059$^{**}$ | 0.062$^{***}$ | 0.058$^{**}$ | 0.061$^{***}$ | 0.057$^{**}$ | 0.061$^{***}$ | 0.052$^{**}$ |
| | | | (0.043) | (0.043) | (0.023) | (0.023) | (0.023) | (0.023) | (0.023) | (0.023) | (0.023) | (0.023) |
| **Distance to Nearest Outlet** | | | | | | | | | -0.00004$^{***}$ | -0.00004$^{***}$ | -0.00004$^{***}$ | -0.00004$^{***}$ |
| | | | | | | | | | (0.00001) | (0.00001) | (0.00001) | (0.00001) |
| **Constant** | 6.248$^{***}$ | 6.332$^{***}$ | 6.048$^{***}$ | 6.049$^{***}$ | 6.291$^{***}$ | 6.369$^{***}$ | 6.293$^{***}$ | 6.372$^{***}$ | 6.329$^{***}$ | 6.374$^{***}$ | 6.335$^{***}$ | 6.416$^{***}$ |
| | (0.574) | (0.575) | (0.027) | (0.027) | (0.573) | (0.575) | (0.573) | (0.574) | (0.107) | (0.107) | (0.085) | (0.085) |
| **Observations** | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 | 3,512 |
| **$R^2$** | 0.747 | 0.746 | 0.041 | 0.041 | 0.748 | 0.747 | 0.749 | 0.747 | 0.749 | 0.748 | 0.749 | 0.747 |
| **Adjusted $R^2$** | 0.745 | 0.744 | 0.040 | 0.040 | 0.746 | 0.745 | 0.746 | 0.745 | 0.747 | 0.746 | 0.748 | 0.746 |
| **Residual Std. Error** | 0.304 (df = 3485) | 0.305 (df = 3485) | 0.589 (df = 3506) | 0.589 (df = 3506) | 0.303 (df = 3480) | 0.304 (df = 3480) | 0.303 (df = 3480) | 0.304 (df = 3480) | 0.302 (df = 3483) | 0.303 (df = 3483) | 0.302 (df = 3492) | 0.303 (df = 3494) |
| **F Statistic** | 395.654$^{***}$ (df = 26; 3485) | 392.766$^{***}$ (df = 26; 3485) | 30.030$^{***}$ (df = 5; 3506) | 30.225$^{***}$ (df = 5; 3506) | 333.711$^{***}$ (df = 31; 3480) | 331.038$^{***}$ (df = 31; 3480) | 334.136$^{***}$ (df = 31; 3480) | 331.436$^{***}$ (df = 31; 3480) | 372.035$^{***}$ (df = 28; 3483) | 368.792$^{***}$ (df = 28; 3483) | 548.452$^{***}$ (df = 19; 3492) | 607.437$^{***}$ (df = 17; 3494) |

*Note:*
Significance: *p<0.1; **p<0.05; ***p<0.01
Model 1 - All socio-economic variables - After Tax Income
Model 2 - All socio-economic variables - Before - Tax Income
Model 3 - All built-environment variables - Outlet Den. by Pop
Model 4 - All built-environment variables - Outlet Den. by Area
Model 5 - All socio-economic and built-environment - After-Tax - Outlet Den by Pop
Model 6 - All socio-economic and built-environment - Before-Tax - Outlet Dent by Pop
Model 7 - All socio-economic and built-environment - After-Tax - Outlet Den by Area
Model 8 - All socio-economic and built-environment - Before -Tax - Outlet Den by Area
Model 9 - All significant variables - After - Tax Income
Model 10 - All significant variables - Before - Tax Income
Model 11 - Backwards Regression - 0.1 alpha - After – Tax
Model 12 - Backwards Regression - 0.1 alpha - Before-Tax

# A7. Survey of Household Spending Questionnaire



FOOD (FO)

FO_R001 The next section refers to your food, alcohol and cigarette purchases.

FO_Q010 In the last 4 weeks (^DT_4W_E), how much did your household spend on:
... food purchased from stores?
Include purchases from grocery stores, specialty food stores, department stores, warehouse-type stores, convenience stores and home delivery.
Do not include non-food items bought with your groceries.

|_|_|_|_|_|_|_|
(MIN: 0) (MAX: 99999997)
DK, RF

? FO_Q020 (In the last 4 weeks, ^DT_4W_E, how much did your household spend on:)
... alcoholic beverages purchased from stores?
Include purchases made in liquor stores, beer stores, wine stores and grocery stores.
Do not include non-alcoholic beer and wine.

|_|_|_|_|_|_|_|
(MIN: 0) (MAX: 99999997)
DK, RF

FO_Q030 (In the last 4 weeks, ^DT_4W_E, how much did your household spend on:)
... alcoholic beverages purchased and consumed in bars, cocktail lounges and restaurants?
Include all taxes and tips.

|_|_|_|_|_|_|_|
(MIN: 0) (MAX: 99999997)
DK, RF

FO_Q040 (In the last 4 weeks, ^DT_4W_E, how much did your household spend on:)
... cigarettes?

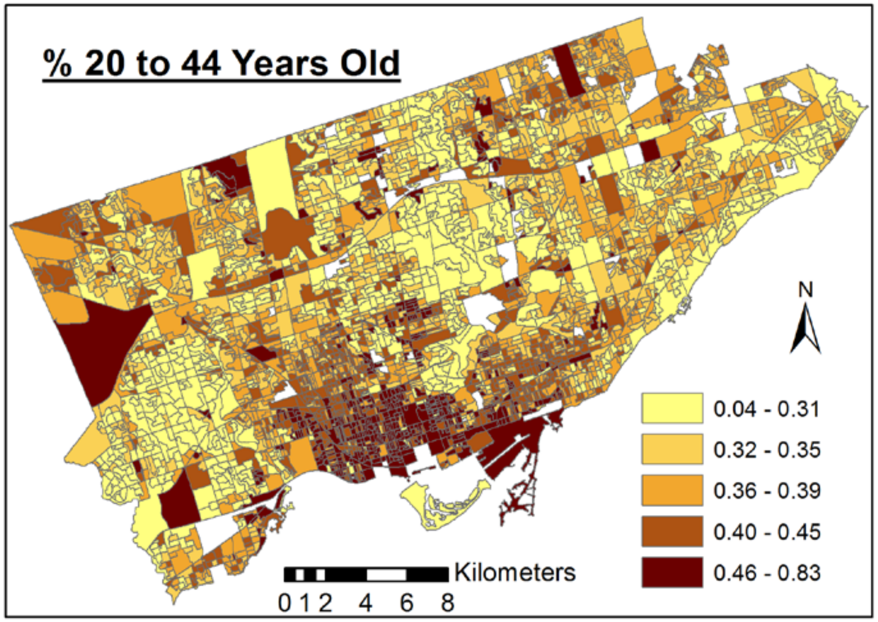
|_|_|_|_|_|_|_|
(MIN: 0) (MAX: 99999997)
DK, RF

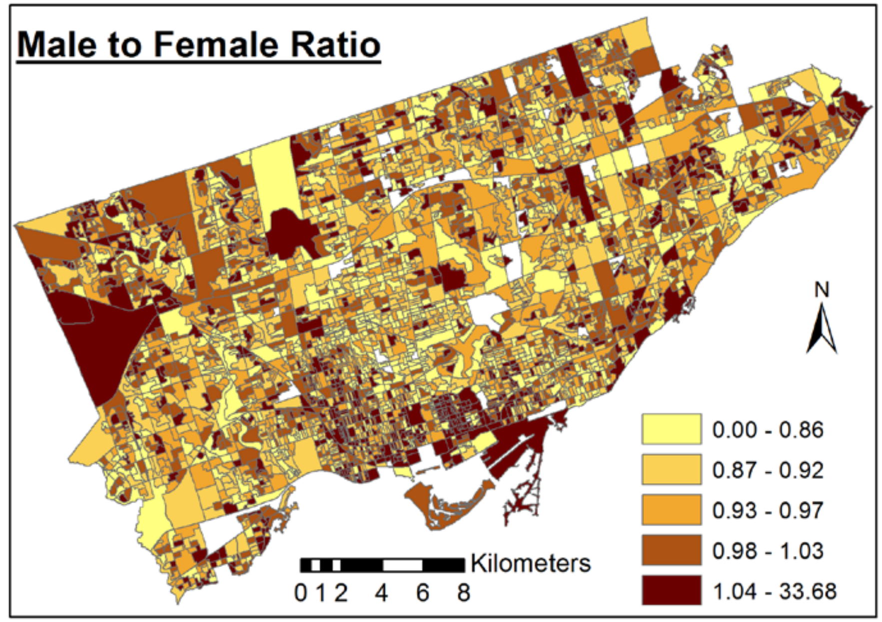
FO_R050 For the next question, please note that the reference period is 12 months.

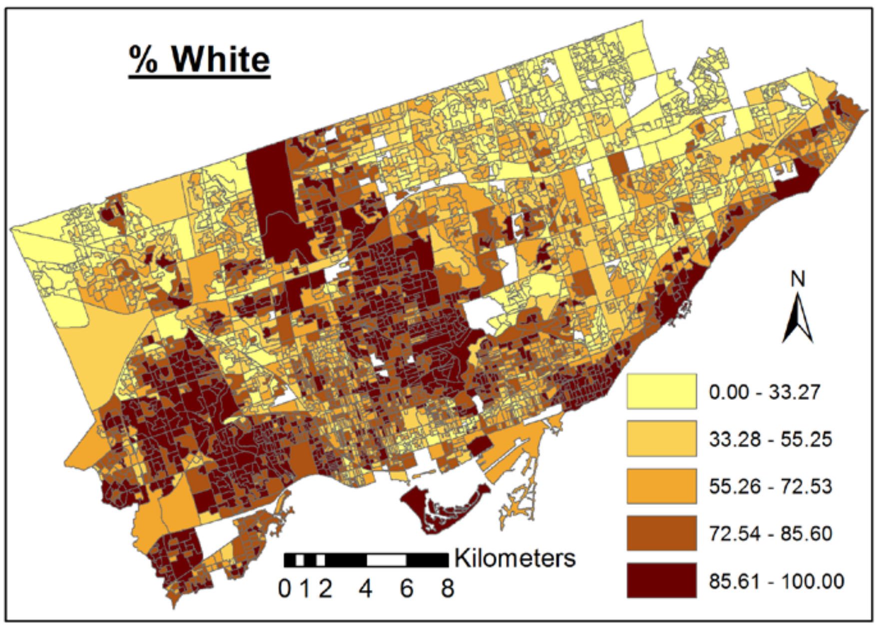
FO_Q050 In the last 12 months (^DT_12M_E), what was the estimated value of any gifts of food, food from your own farm or garden, or from hunting or fishing that added to your supplies?

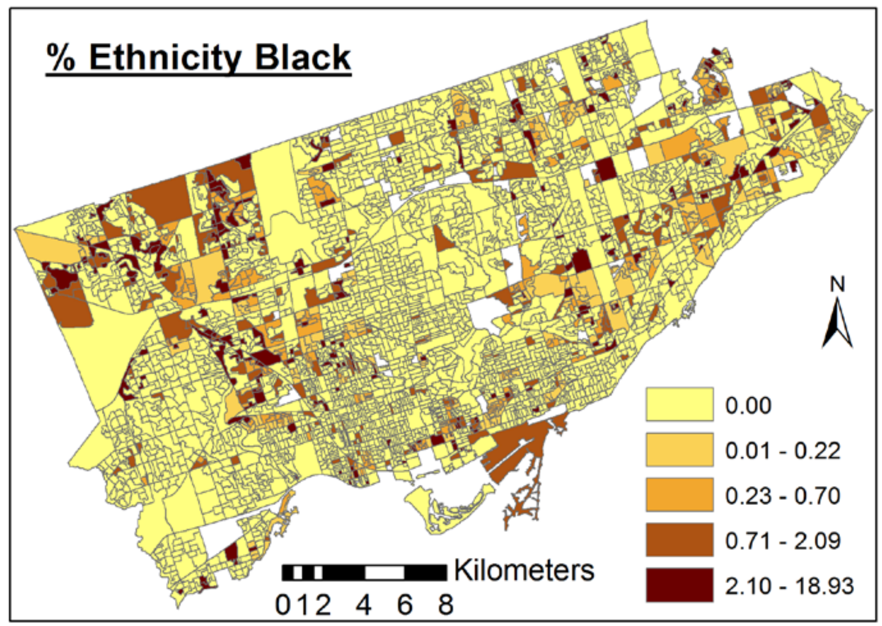
|_|_|_|_|_|_|_|
(MIN: 0) (MAX: 99999997)
DK, RF

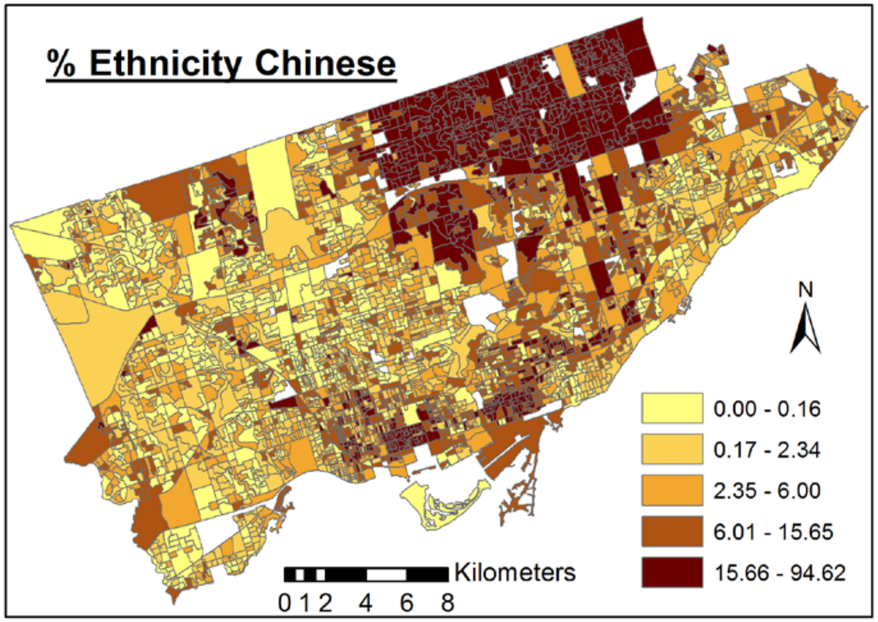
## A8. Visualizations: Quantile Choropleth Maps for All Explanatory Variables

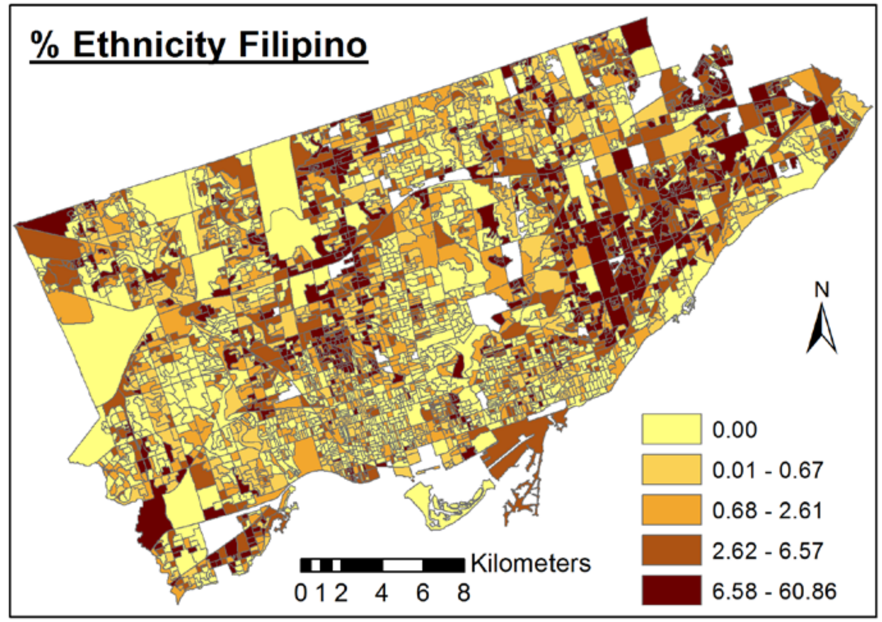

% 20 to 44 Years Old
N
0.04 - 0.31
0.32 - 0.35
0.36 - 0.39
0.40 - 0.45
0.46 - 0.83
Kilometers
0 1 2 4 6 8

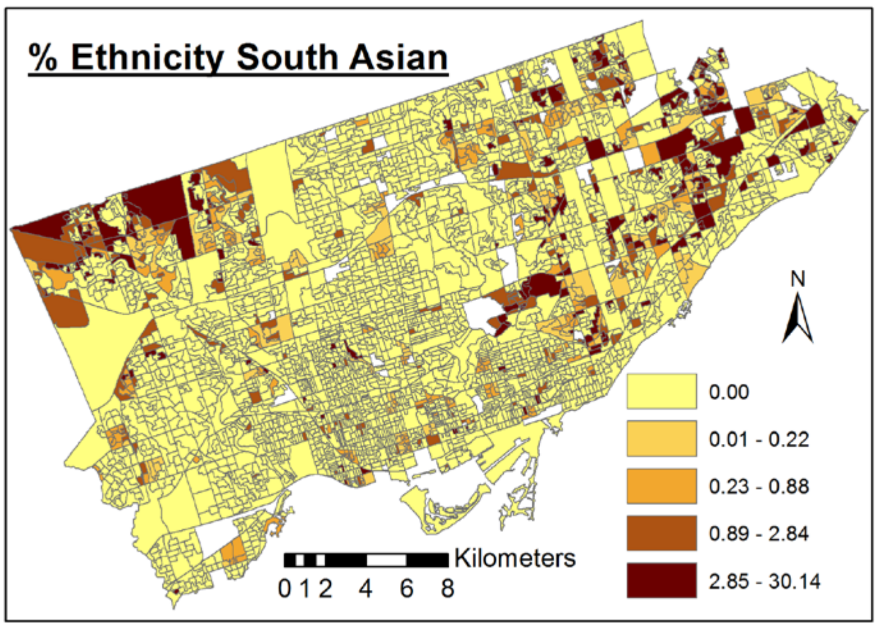

Male to Female Ratio
N
0.00 - 0.86
0.87 - 0.92
0.93 - 0.97
0.98 - 1.03
1.04 - 33.68
Kilometers
0 1 2 4 6 8

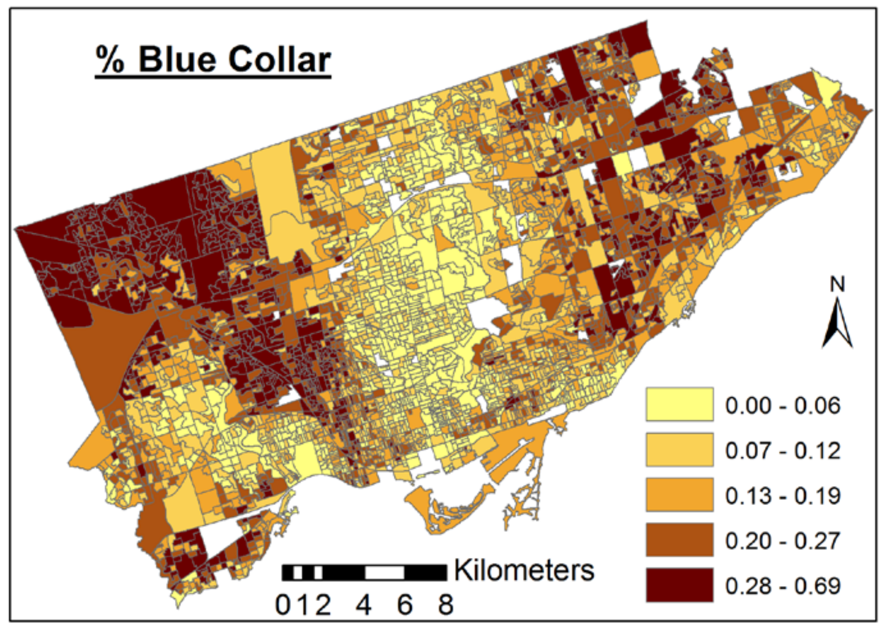

% White
N
0.00 - 33.27
33.28 - 55.25
55.26 - 72.53
72.54 - 85.60
85.61 - 100.00
Kilometers
0 1 2 4 6 8

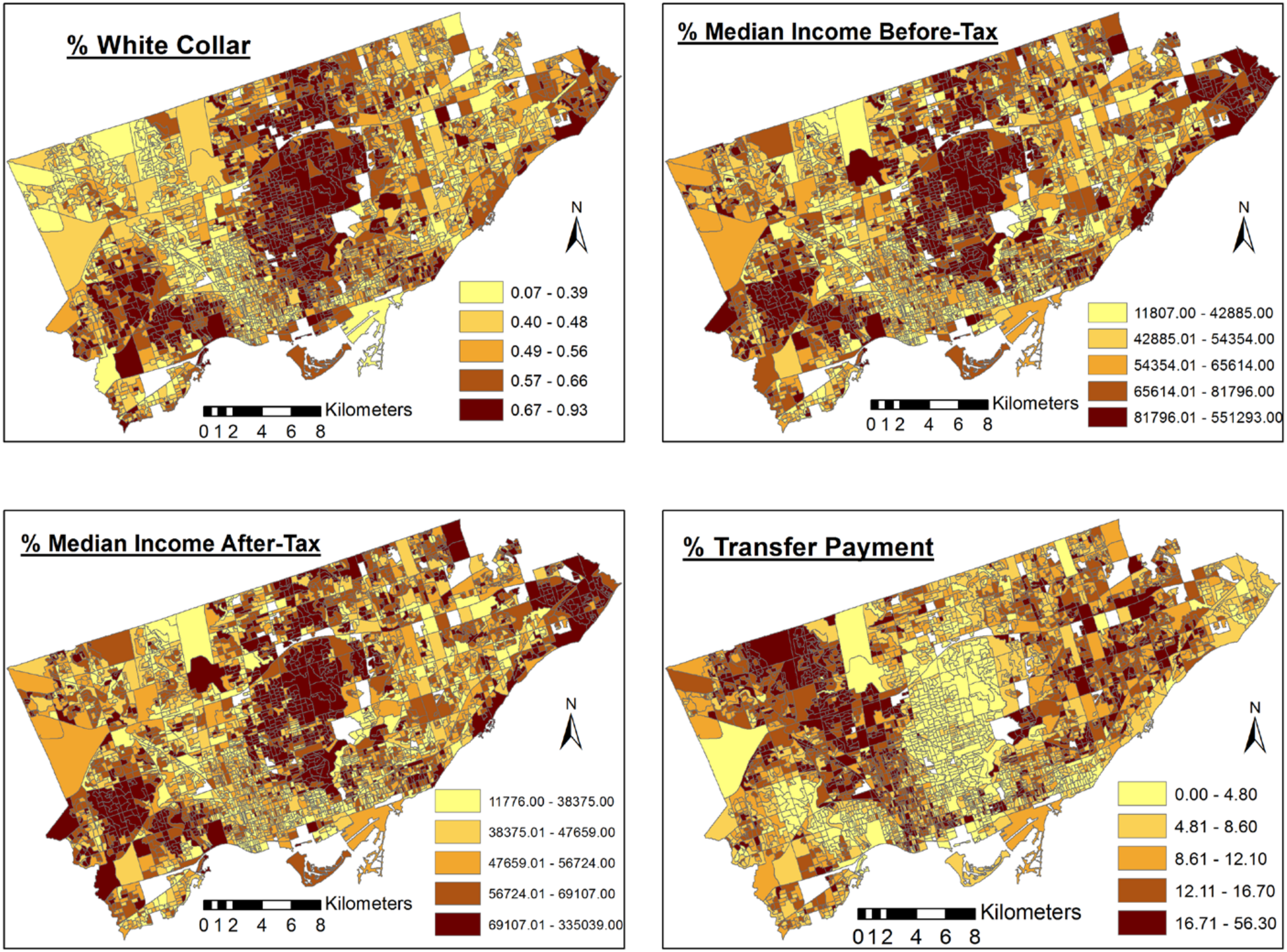

% Ethnicity Black
N
0.00
0.01 - 0.22
0.23 - 0.70
0.71 - 2.09
2.10 - 18.93
Kilometers
0 1 2 4 6 8

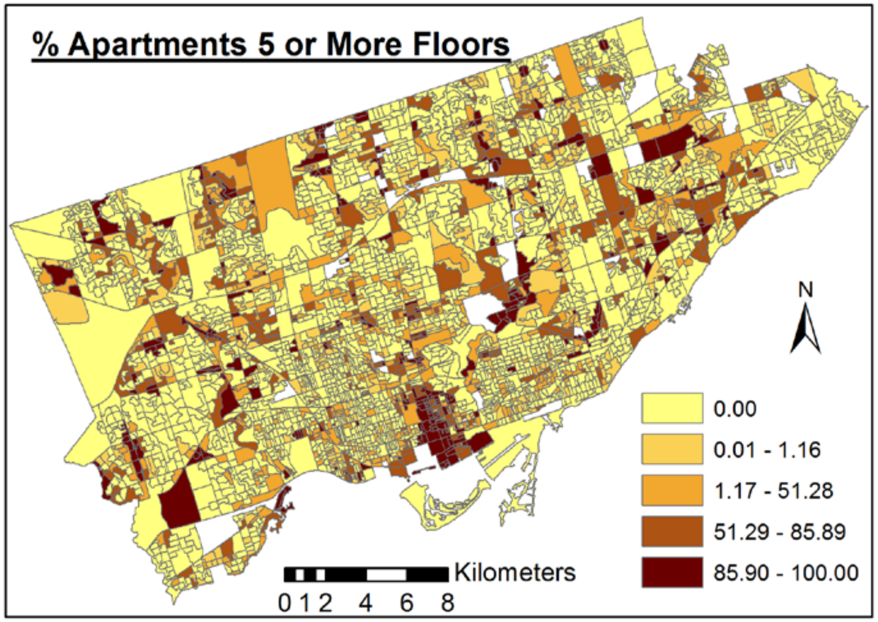

% Ethnicity Chinese
N
0.00 - 0.16
0.17 - 2.34
2.35 - 6.00
6.01 - 15.65
15.66 - 94.62
Kilometers
0 1 2 4 6 8

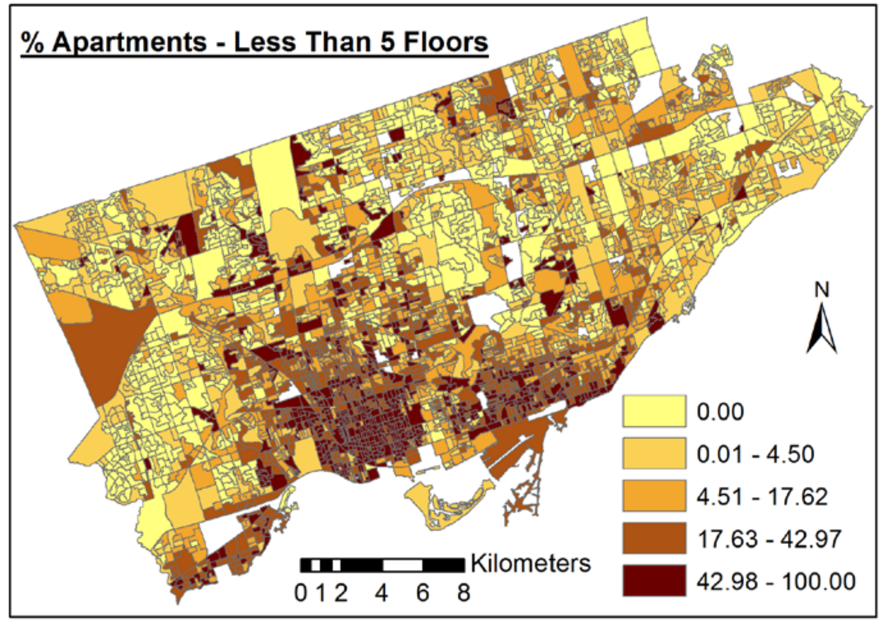

% Ethnicity Filipino
N
0.00
0.01 - 0.67
0.68 - 2.61
2.62 - 6.57
6.58 - 60.86
Kilometers
0 1 2 4 6 8

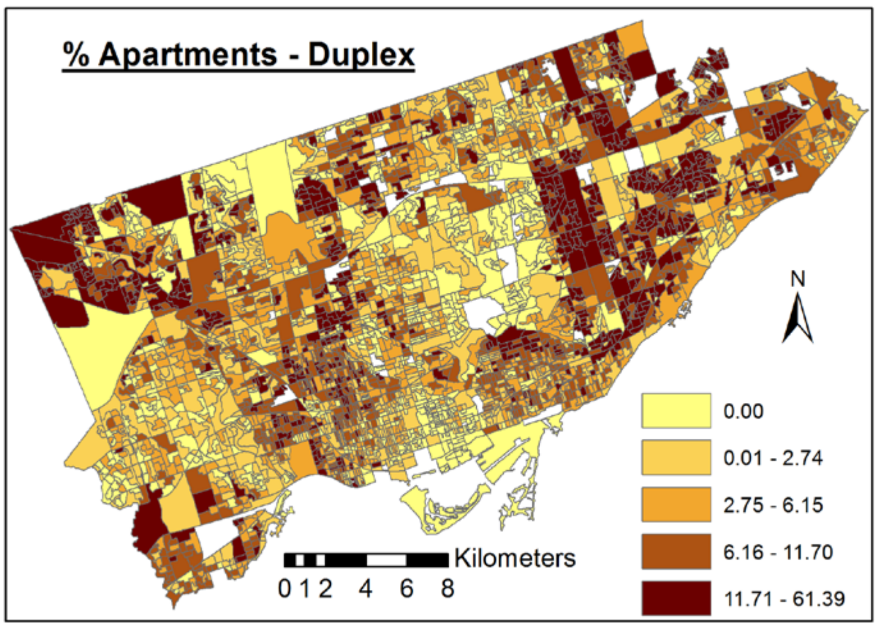

% Ethnicity South Asian
N
0.00
0.01 - 0.22
0.23 - 0.88
0.89 - 2.84
2.85 - 30.14
Kilometers
0 1 2 4 6 8

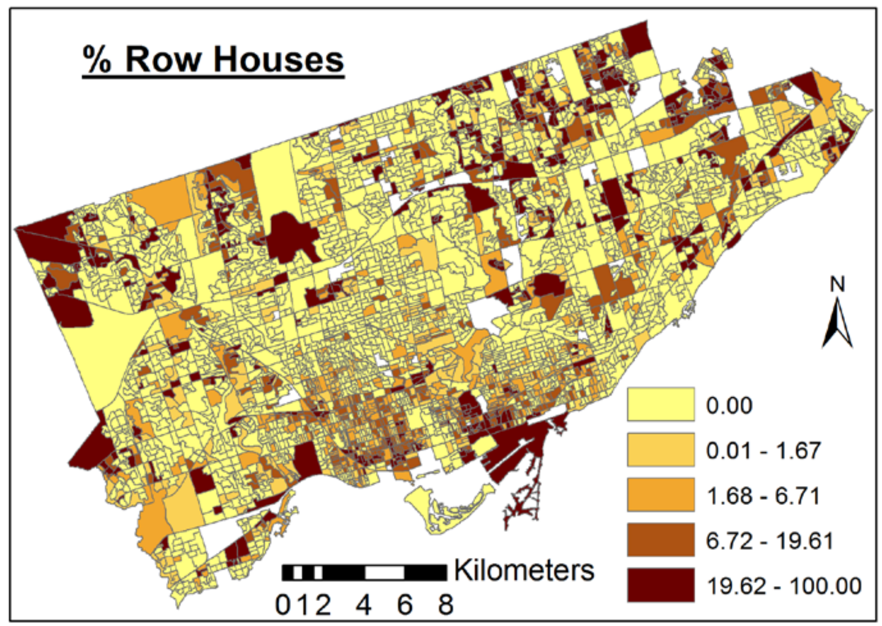

% Blue Collar
N
0.00 - 0.06
0.07 - 0.12
0.13 - 0.19
0.20 - 0.27
0.28 - 0.69
Kilometers
0 1 2 4 6 8

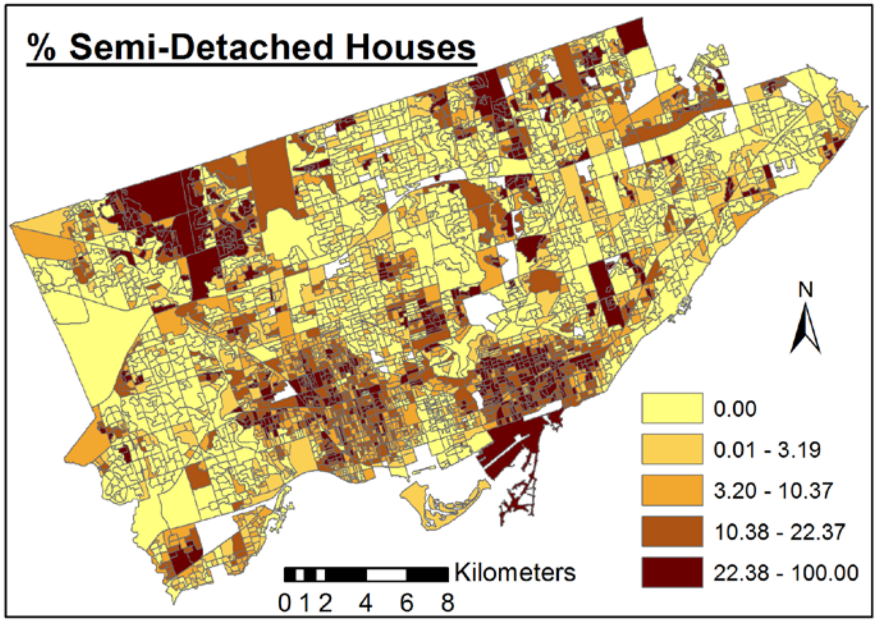

% White Collar
N
0.07 - 0.39
0.40 - 0.48
0.49 - 0.56
0.57 - 0.66
0.67 - 0.93
Kilometers
0 1 2 4 6 8
% Median Income Before-Tax
N
11807.00 - 42885.00
42885.01 - 54354.00
54354.01 - 65614.00
65614.01 - 81796.00
81796.01 - 551293.00
Kilometers
0 1 2 4 6 8
% Median Income After-Tax
N
11776.00 - 38375.00
38375.01 - 47659.00
47659.01 - 56724.00
56724.01 - 69107.00
69107.01 - 335039.00
Kilometers
0 1 2 4 6 8
% Transfer Payment
N
0.00 - 4.80
4.81 - 8.60
8.61 - 12.10
12.11 - 16.70
16.71 - 56.30
Kilometers
0 1 2 4 6 8


% Apartments 5 or More Floors
N
0.00
0.01 - 1.16
1.17 - 51.28
51.29 - 85.89
85.90 - 100.00
Kilometers
0 1 2 4 6 8

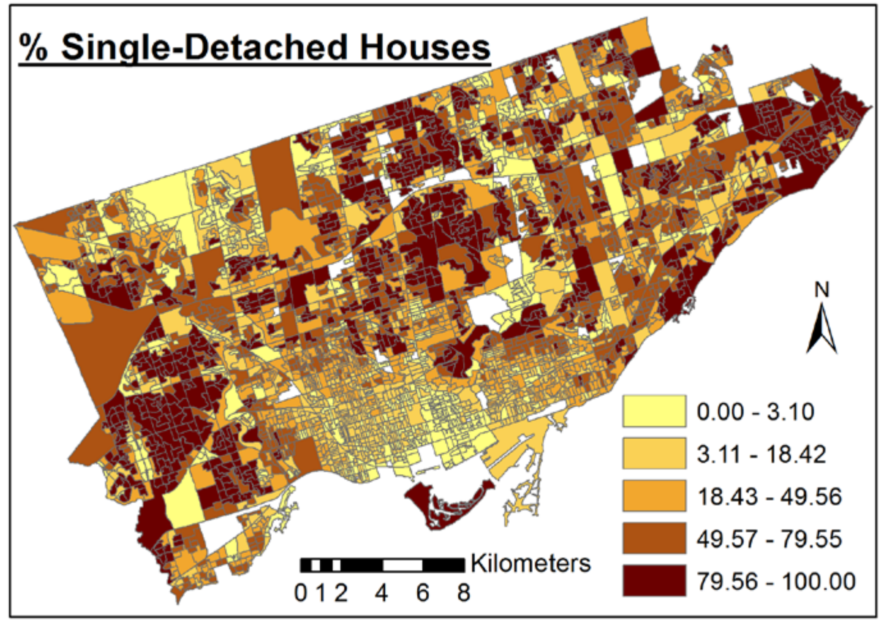

% Apartments - Less Than 5 Floors
N
0.00
0.01 - 4.50
4.51 - 17.62
17.63 - 42.97
42.98 - 100.00
Kilometers
0 1 2 4 6 8


% Apartments - Duplex
N
0.00
0.01 - 2.74
2.75 - 6.15
6.16 - 11.70
11.71 - 61.39
Kilometers
0 1 2 4 6 8

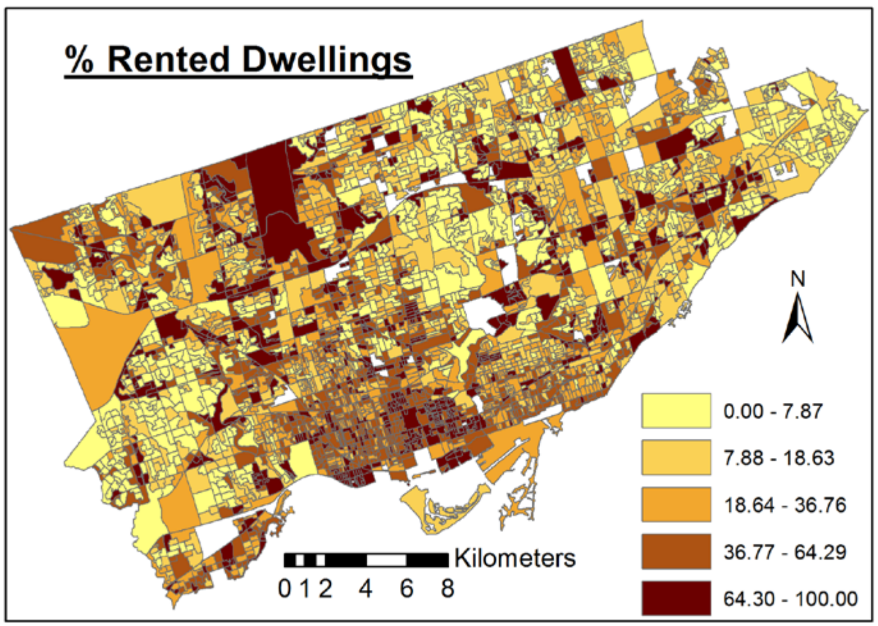

% Row Houses
N
0.00
0.01 - 1.67
1.68 - 6.71
6.72 - 19.61
19.62 - 100.00
Kilometers
0 1 2 4 6 8



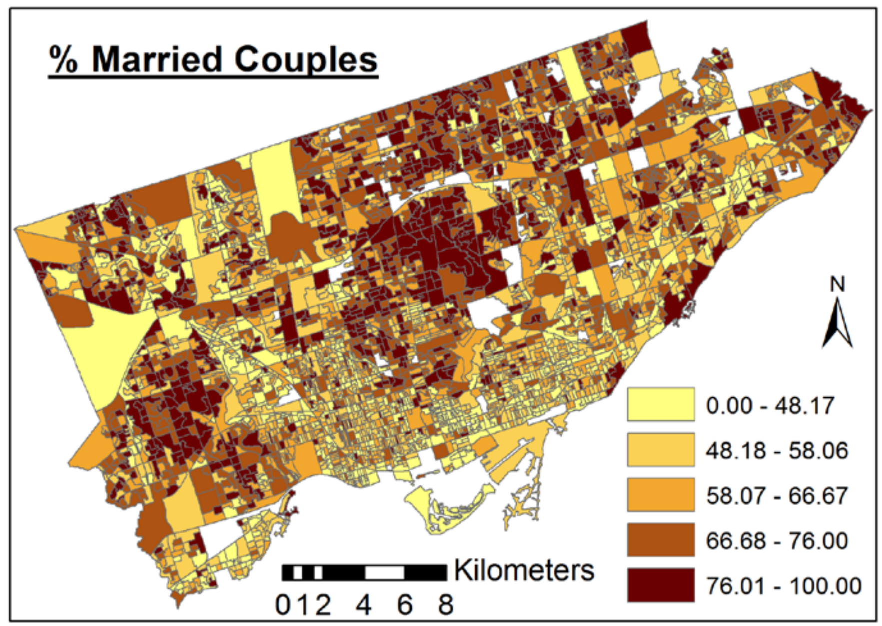
% Semi-Detached Houses
N
0.00
0.01 - 3.19
3.20 - 10.37
10.38 - 22.37
22.38 - 100.00
Kilometers
0 1 2 4 6 8

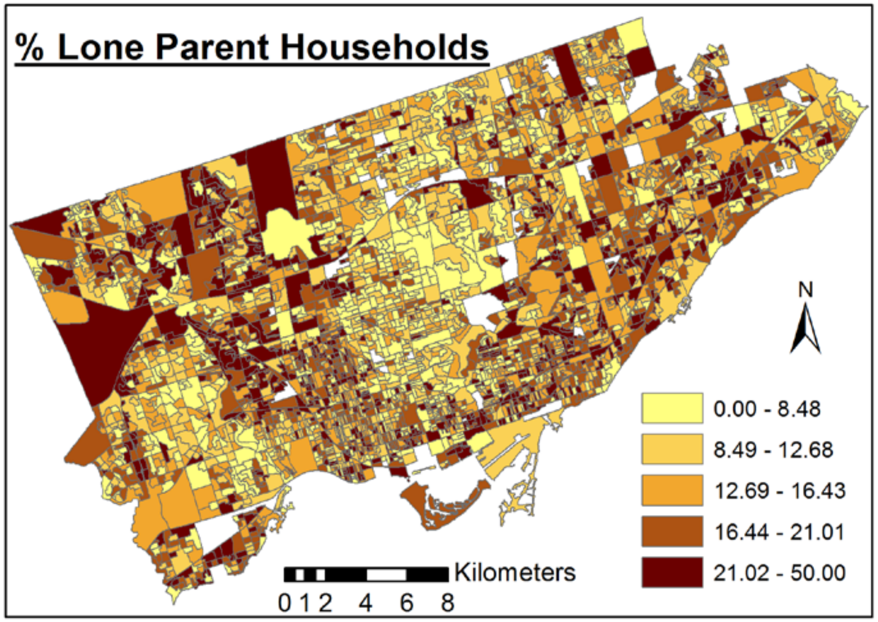
% Single-Detached Houses
N
0.00 - 3.10
3.11 - 18.42
18.43 - 49.56
49.57 - 79.55
79.56 - 100.00
Kilometers
0 1 2 4 6 8

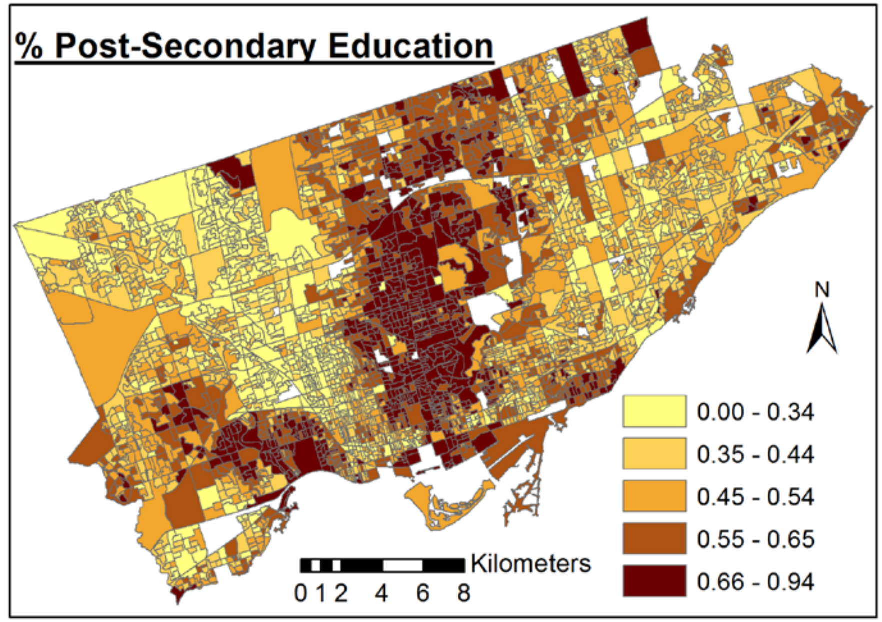



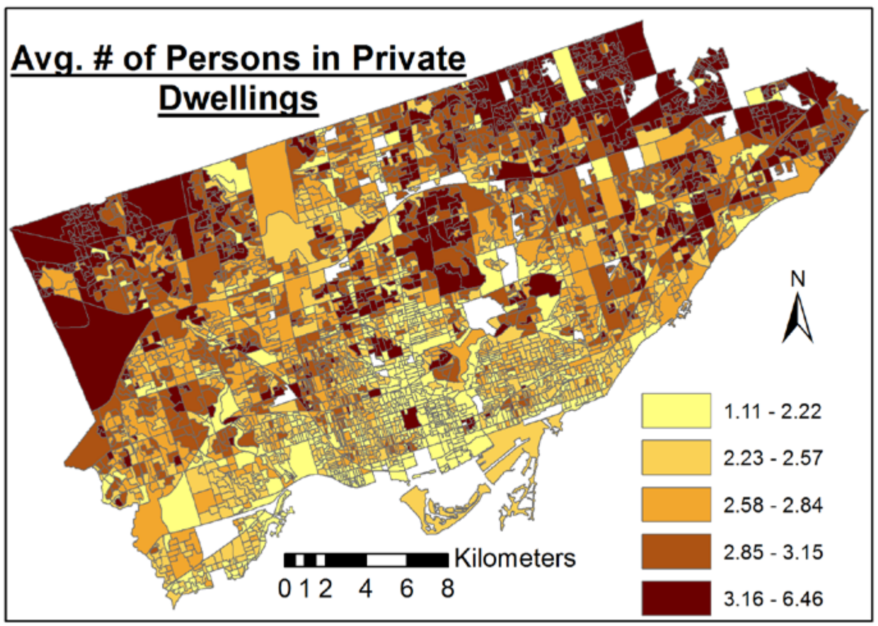
% Rented Dwellings
N
0.00 - 7.87
7.88 - 18.63
18.64 - 36.76
36.77 - 64.29
64.30 - 100.00
Kilometers
0 1 2 4 6 8

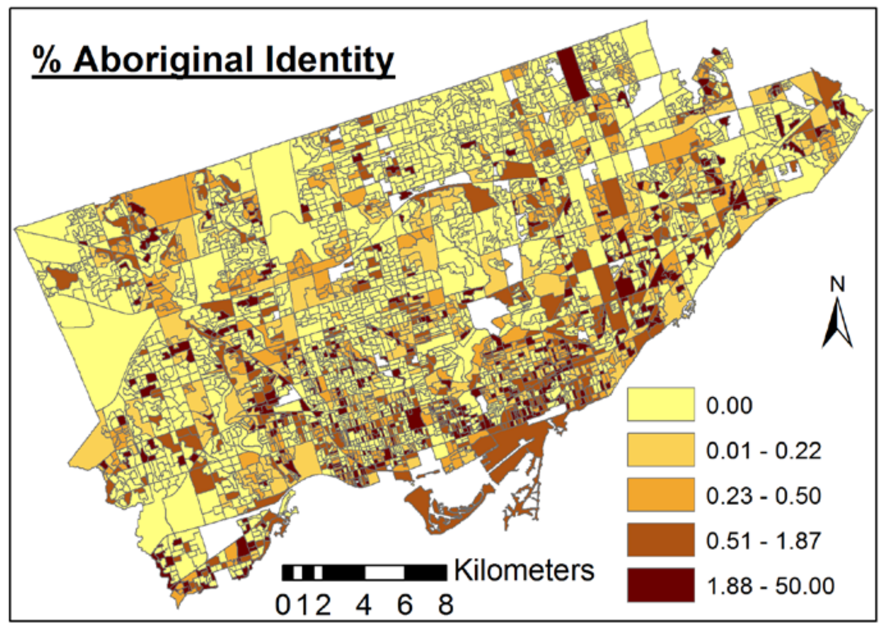



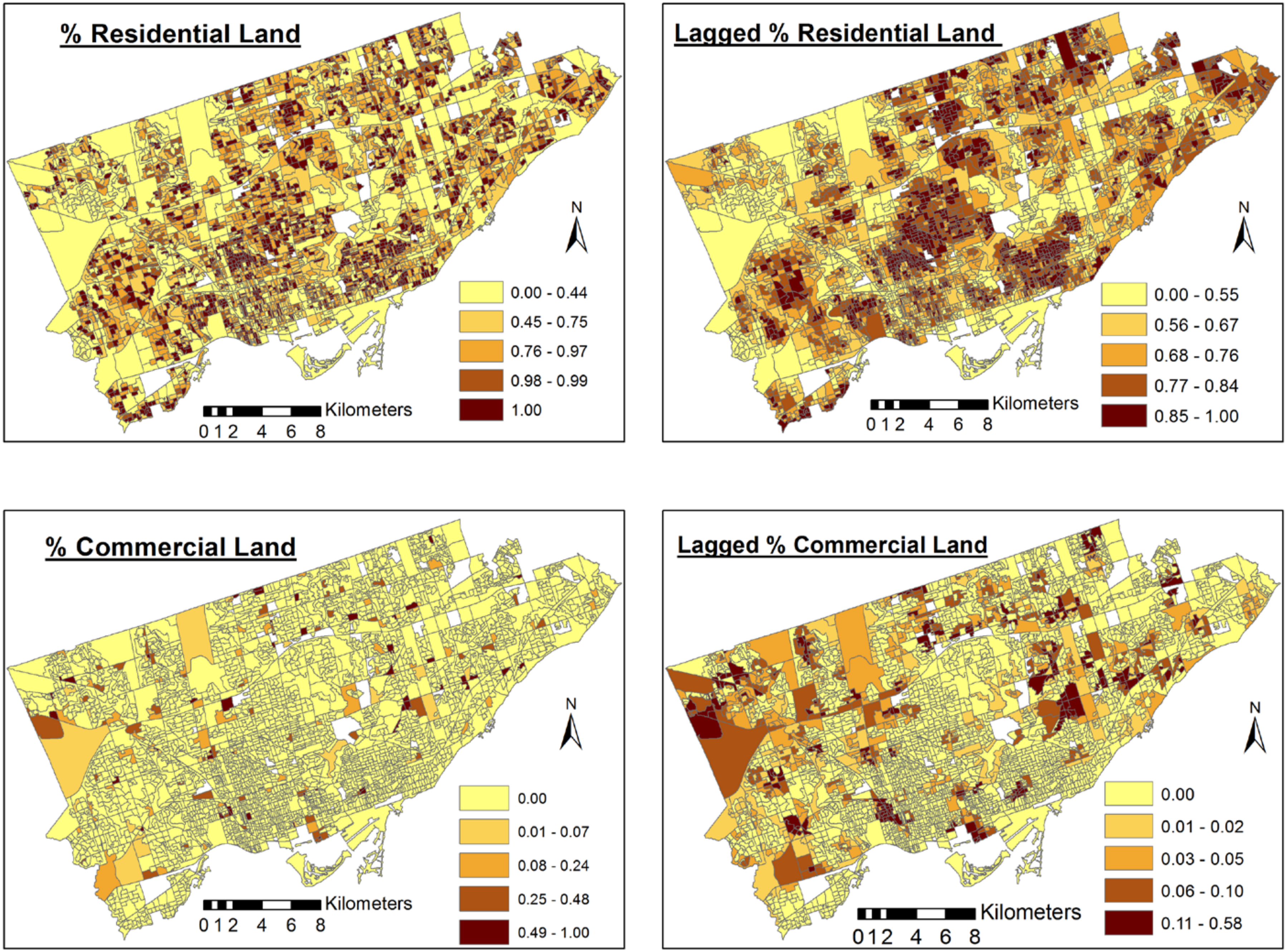
% Married Couples
N
0.00 - 48.17
48.18 - 58.06
58.07 - 66.67
66.68 - 76.00
76.01 - 100.00
Kilometers
0 1 2 4 6 8

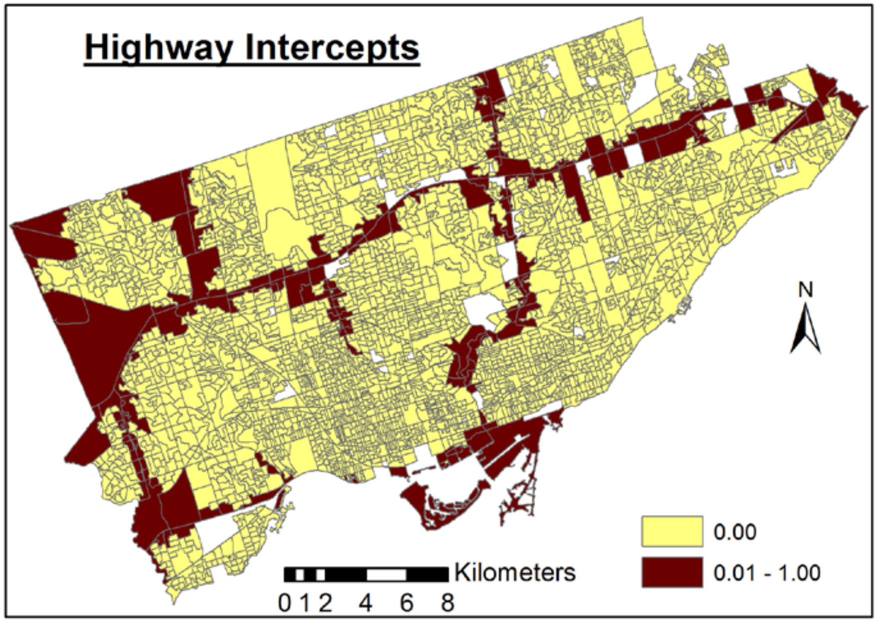

% Lone Parent Households
N
0.00 - 8.48
8.49 - 12.68
12.69 - 16.43
16.44 - 21.01
21.02 - 50.00
Kilometers
0 1 2 4 6 8

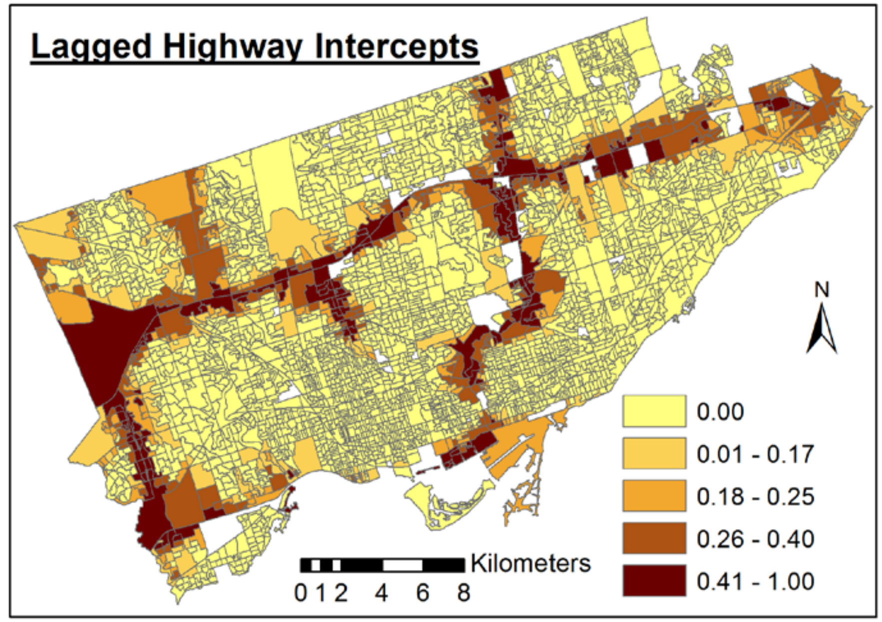

% Post-Secondary Education
N
0.00 - 0.34
0.35 - 0.44
0.45 - 0.54
0.55 - 0.65
0.66 - 0.94
Kilometers
0 1 2 4 6 8

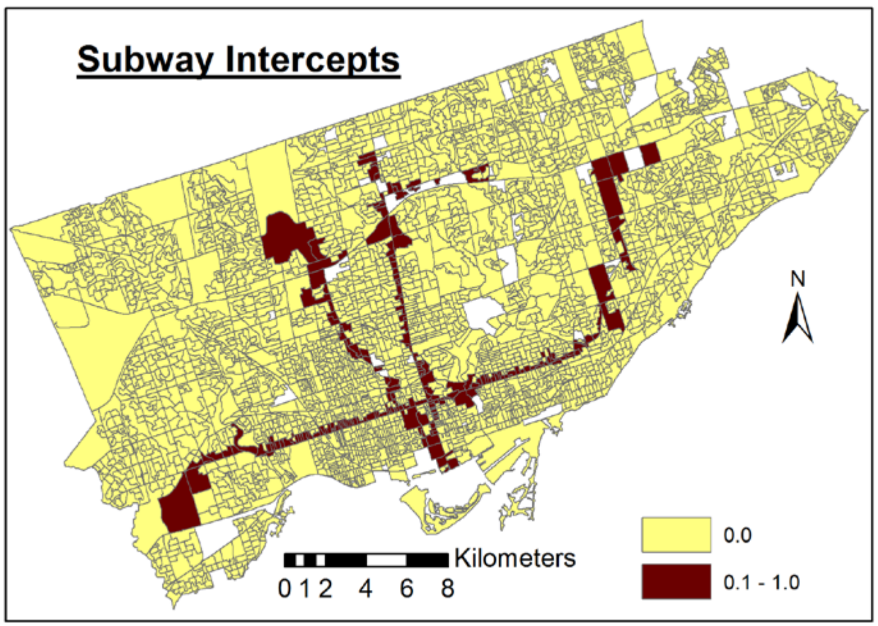

Avg. # of Persons in Private Dwellings
N
1.11 - 2.22
2.23 - 2.57
2.58 - 2.84
2.85 - 3.15
3.16 - 6.46
Kilometers
0 1 2 4 6 8

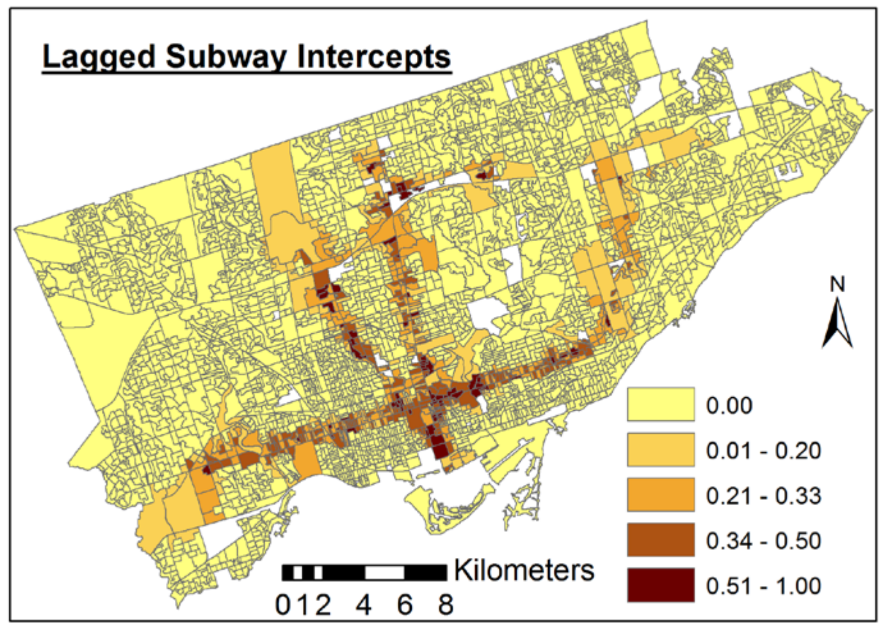

% Aboriginal Identity
N
0.00
0.01 - 0.22
0.23 - 0.50
0.51 - 1.87
1.88 - 50.00
Kilometers
0 1 2 4 6 8

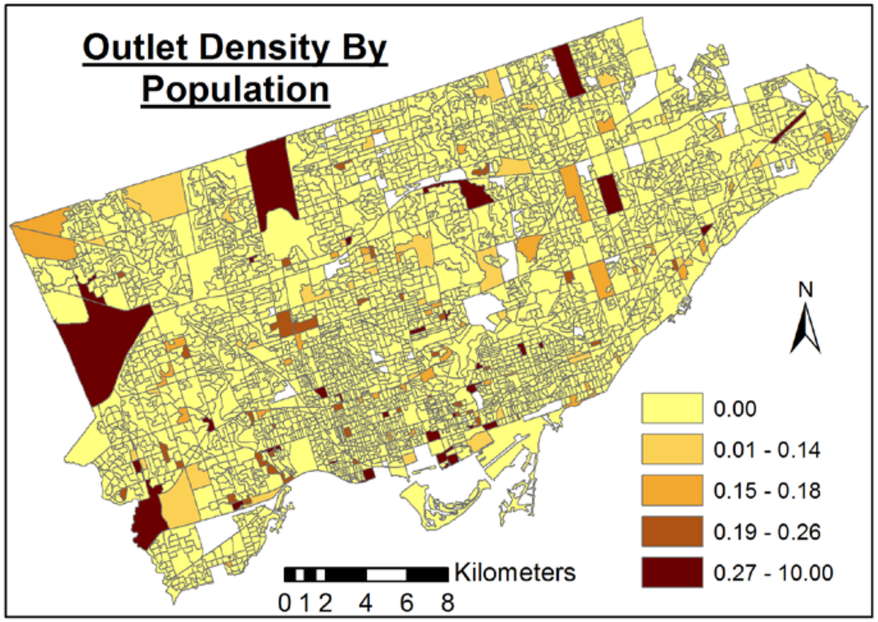

% Residential Land
N
0.00 - 0.44
0.45 - 0.75
0.76 - 0.97
0.98 - 0.99
1.00
Kilometers
0 1 2 4 6 8
Lagged % Residential Land
N
0.00 - 0.55
0.56 - 0.67
0.68 - 0.76
0.77 - 0.84
0.85 - 1.00
Kilometers
0 1 2 4 6 8
% Commercial Land
N
0.00
0.01 - 0.07
0.08 - 0.24
0.25 - 0.48
0.49 - 1.00
Kilometers
0 1 2 4 6 8
Lagged % Commercial Land
N
0.00
0.01 - 0.02
0.03 - 0.05
0.06 - 0.10
0.11 - 0.58
Kilometers
0 1 2 4 6 8

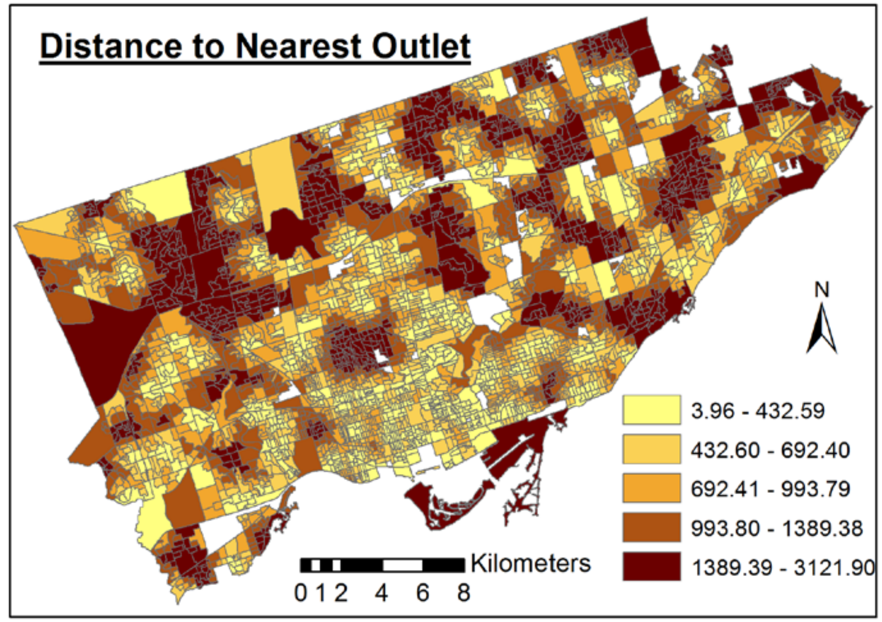

Highway Intercepts
N
Kilometers
0 1 2 4 6 8
0.00
0.01 - 1.00

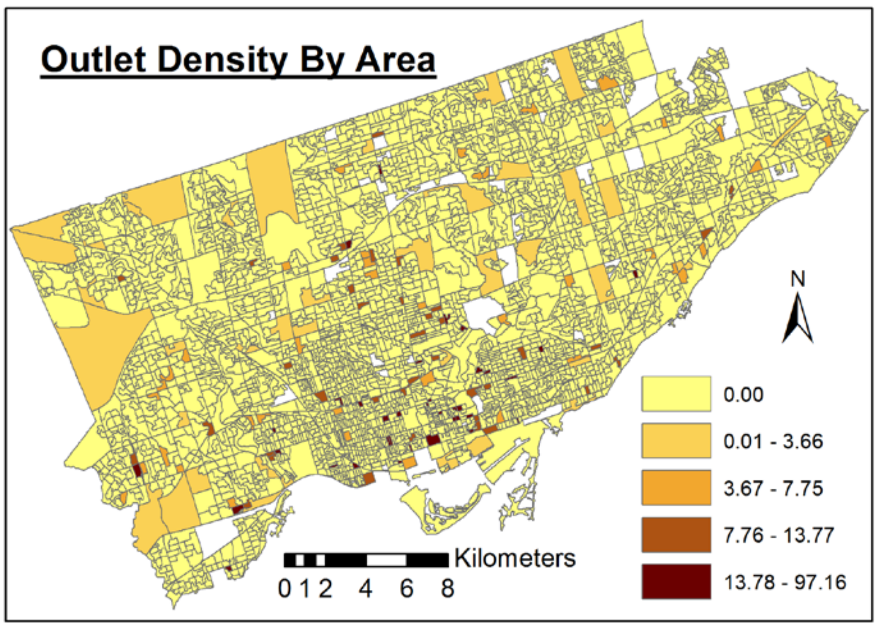

Lagged Highway Intercepts
N
0.00
0.01 - 0.17
0.18 - 0.25
0.26 - 0.40
0.41 - 1.00
Kilometers
0 1 2 4 6 8

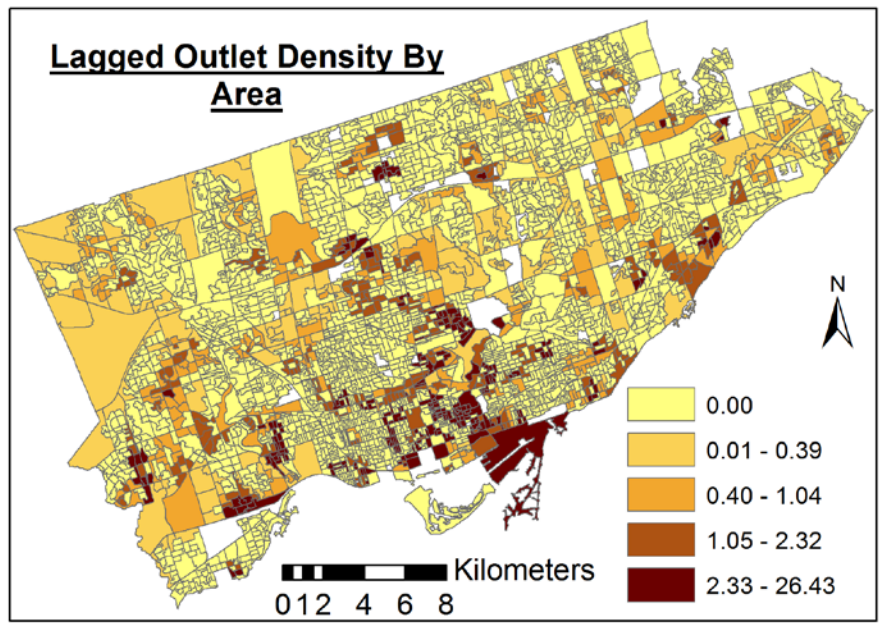

Subway Intercepts
N
Kilometers
0 1 2 4 6 8
0.0
0.1 - 1.0

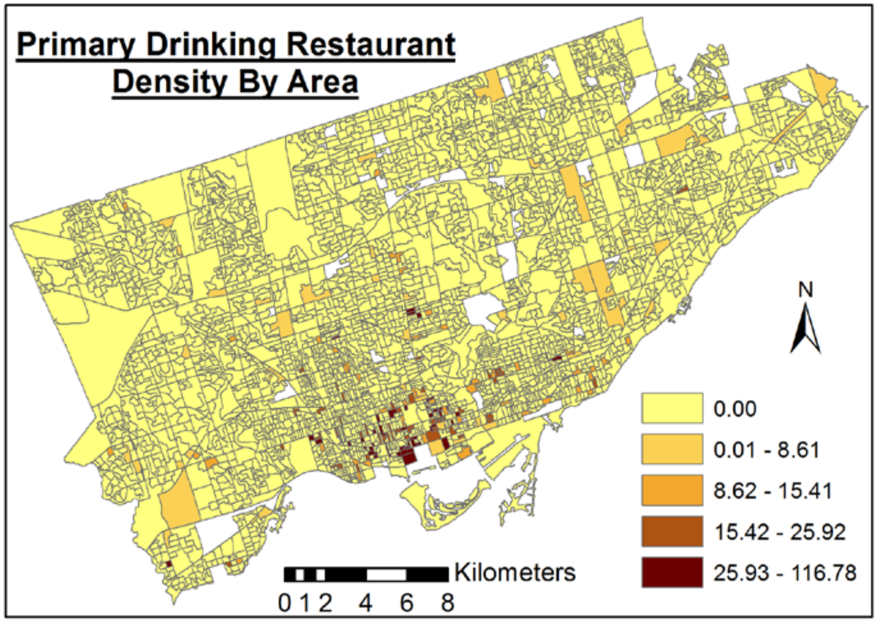

Lagged Subway Intercepts
N
0.00
0.01 - 0.20
0.21 - 0.33
0.34 - 0.50
0.51 - 1.00
Kilometers
0 1 2 4 6 8

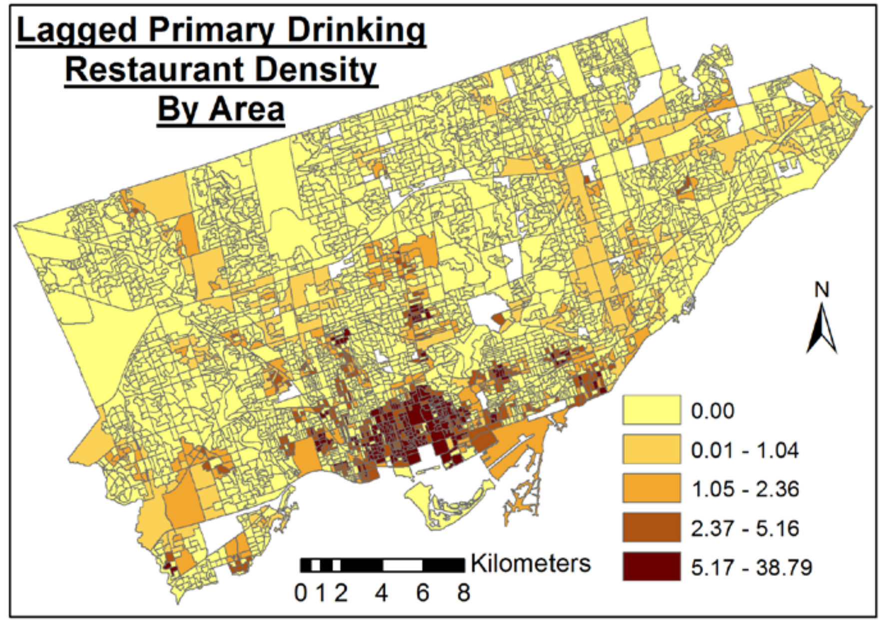

Outlet Density By Population
N
0.00
0.01 - 0.14
0.15 - 0.18
0.19 - 0.26
0.27 - 10.00
Kilometers
0 1 2 4 6 8

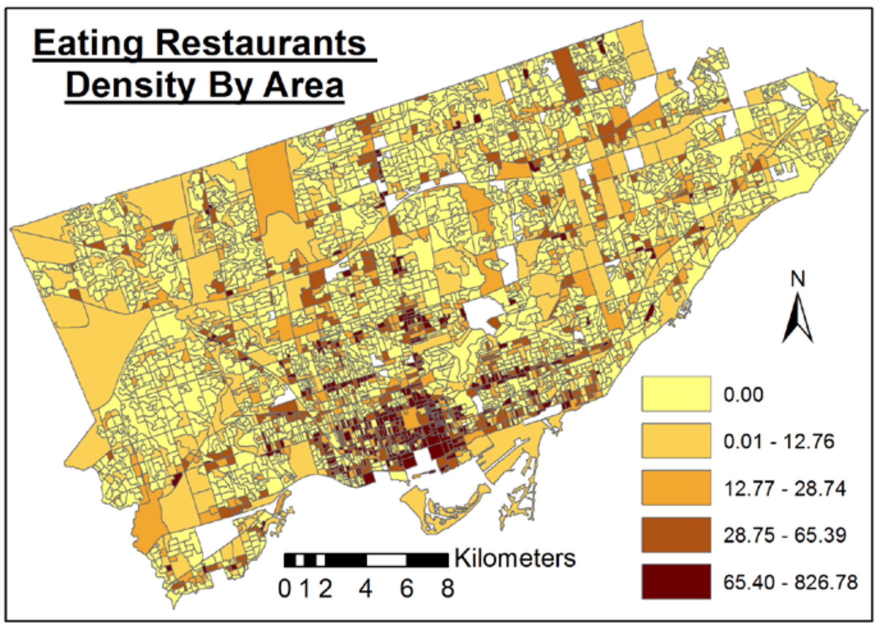

Distance to Nearest Outlet
N
3.96 - 432.59
432.60 - 692.40
692.41 - 993.79
993.80 - 1389.38
1389.39 - 3121.90
Kilometers
0 1 2 4 6 8

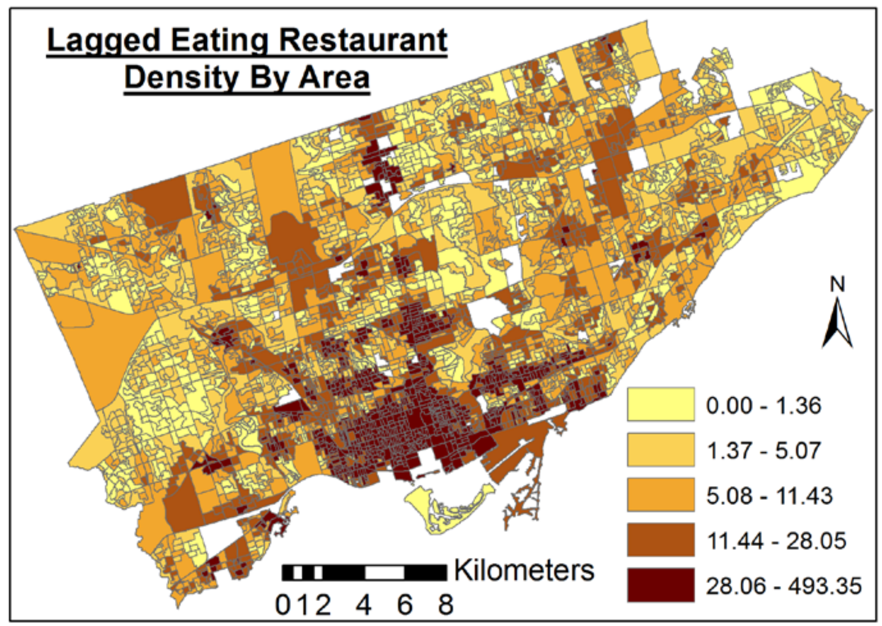

Outlet Density By Area
N
0.00
0.01 - 3.66
3.67 - 7.75
7.76 - 13.77
13.78 - 97.16
Kilometers
0 1 2 4 6 8


Lagged Outlet Density By Area
N
0.00
0.01 - 0.39
0.40 - 1.04
1.05 - 2.32
2.33 - 26.43
Kilometers
0 1 2 4 6 8


Primary Drinking Restaurant Density By Area
0.00
0.01 - 8.61
8.62 - 15.41
15.42 - 25.92
25.93 - 116.78
N
Kilometers
0 1 2 4 6 8

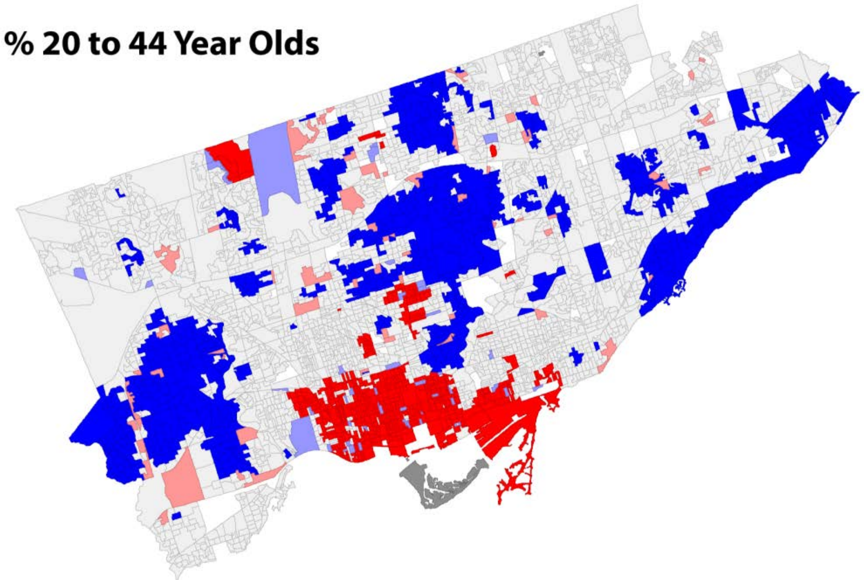

Lagged Primary Drinking Restaurant Density By Area
0.00
0.01 - 1.04
1.05 - 2.36
2.37 - 5.16
5.17 - 38.79
N
Kilometers
0 1 2 4 6 8


Eating Restaurants Density By Area
0.00
0.01 - 12.76
12.77 - 28.74
28.75 - 65.39
65.40 - 826.78
N
Kilometers
0 1 2 4 6 8

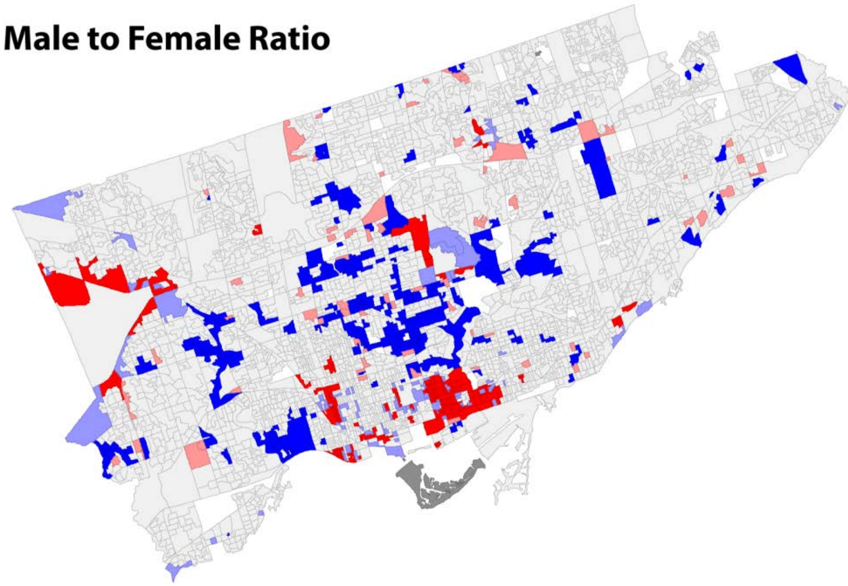

Lagged Eating Restaurant Density By Area
0.00 - 1.36
1.37 - 5.07
5.08 - 11.43
11.44 - 28.05
28.06 - 493.35
N
Kilometers
0 1 2 4 6 8

## A9. Visualizations: Hot-spot and Cold-spot Cluster Maps for All Explanatory Variables

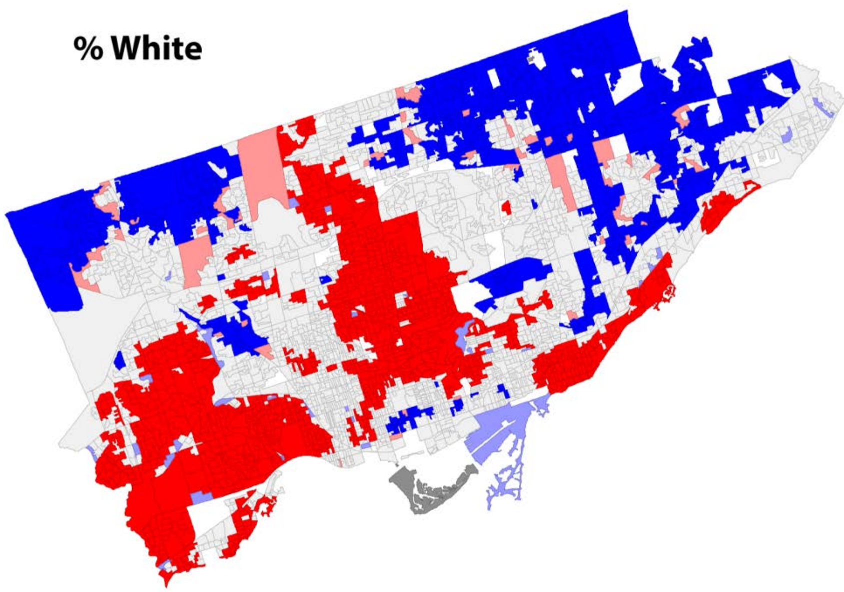
**% 20 to 44 Year Olds**



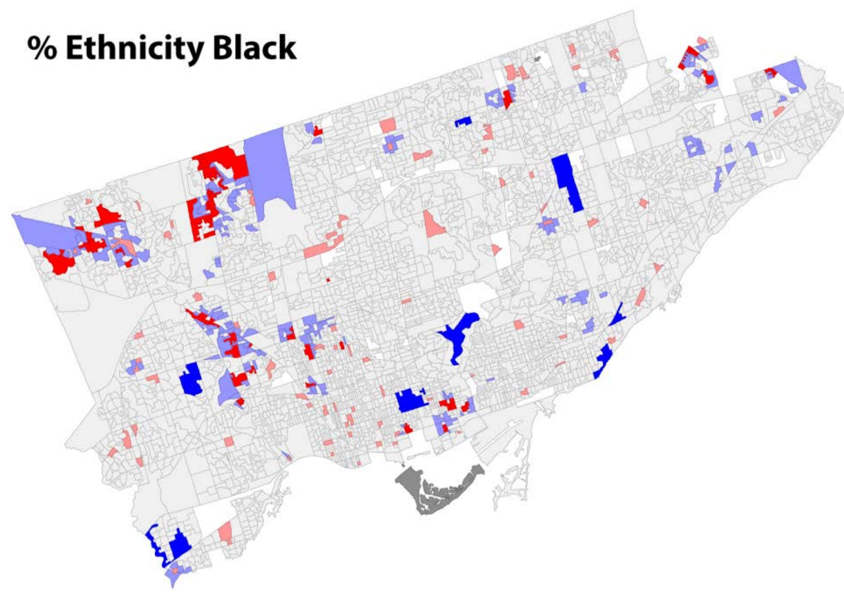
**Male to Female Ratio**



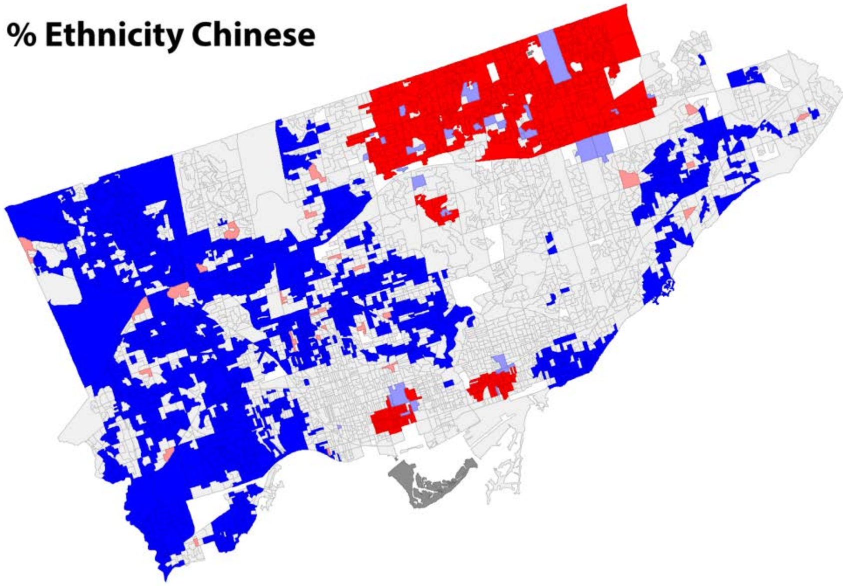
**% White**



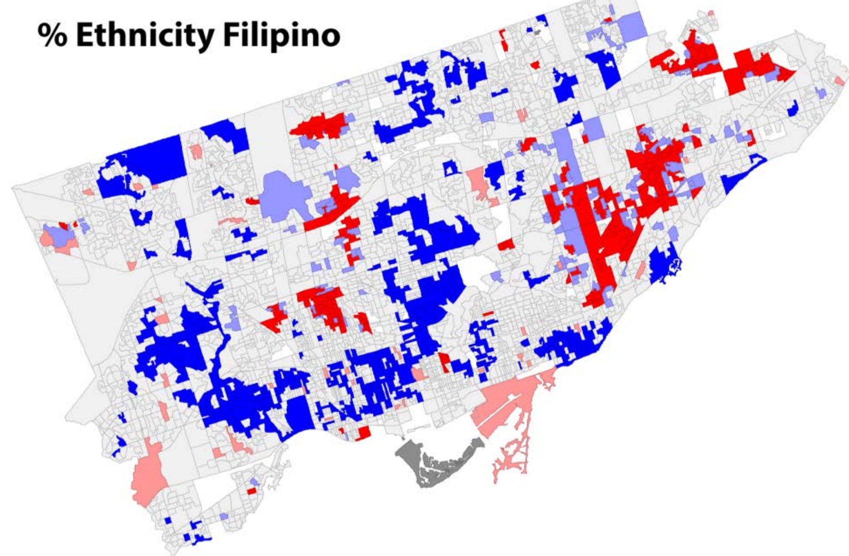
**% Ethnicity Black**



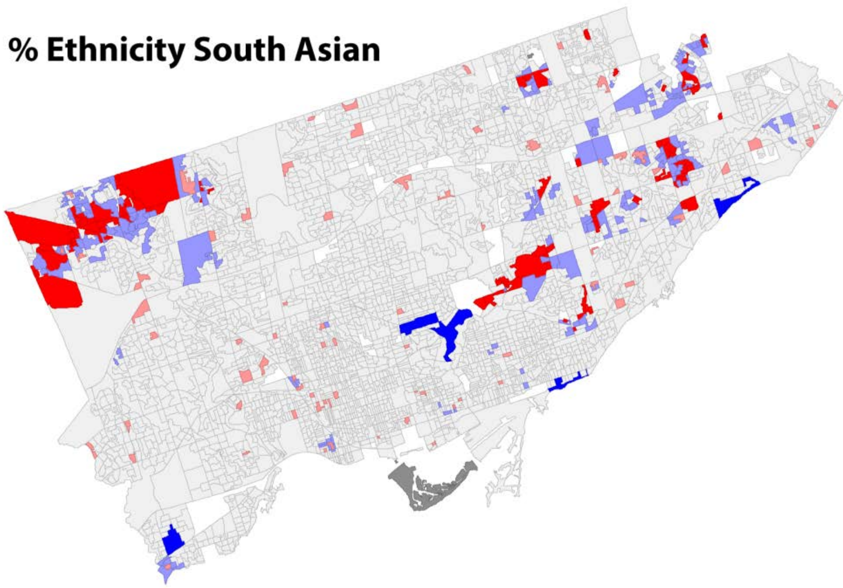
**% Ethnicity Chinese**



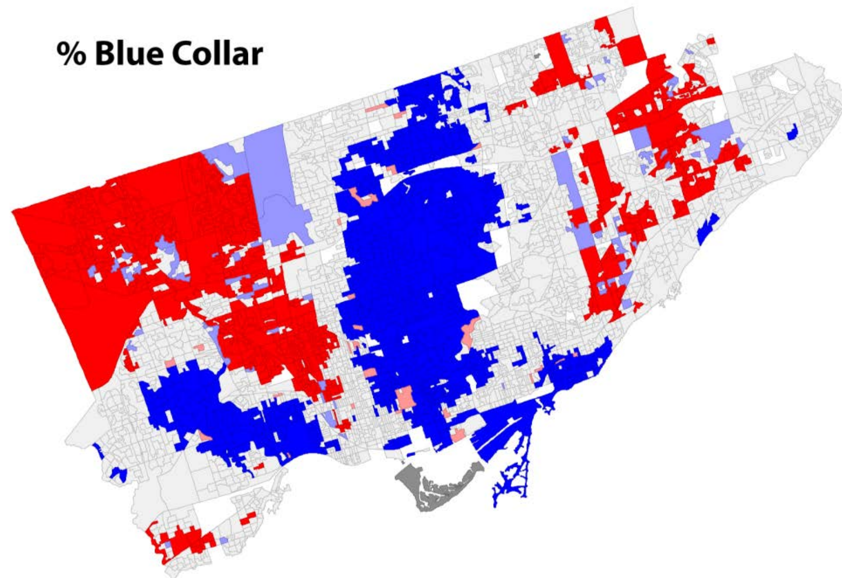
**% Ethnicity Filipino**



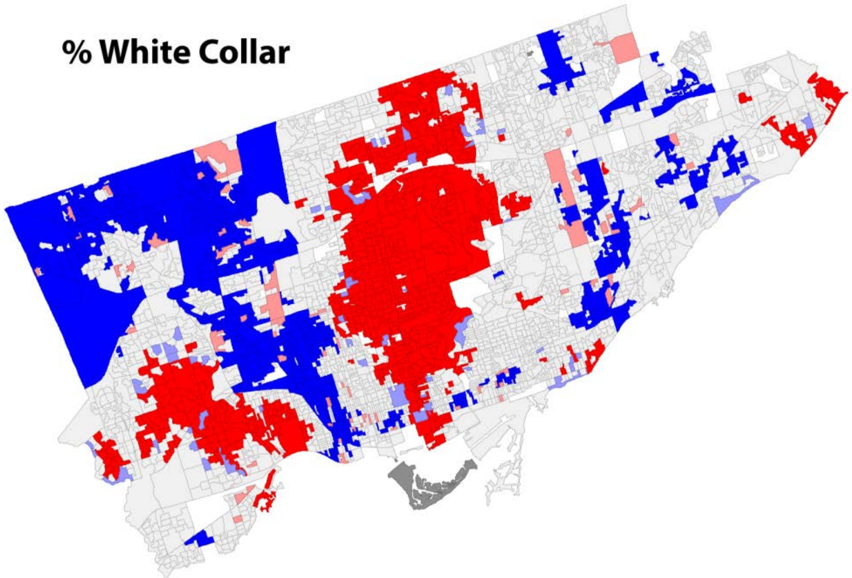
**% Ethnicity South Asian**



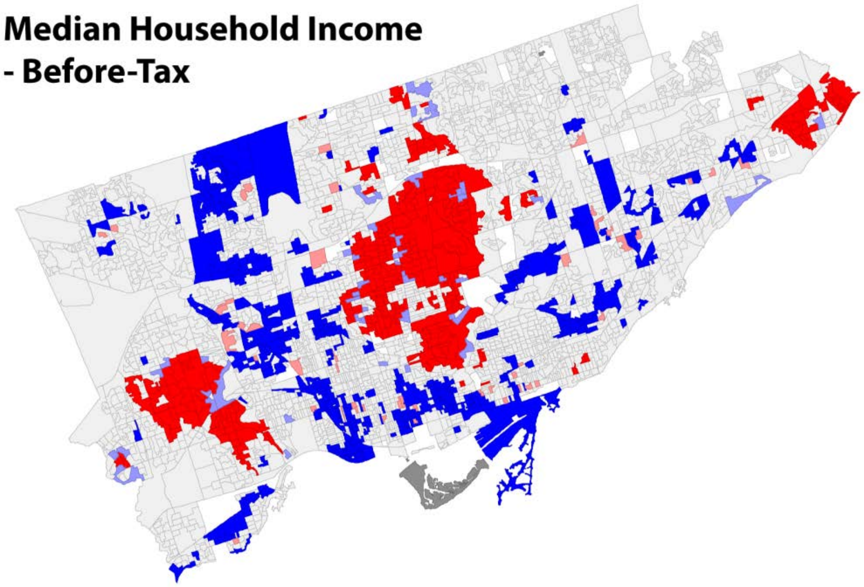
**% Blue Collar**



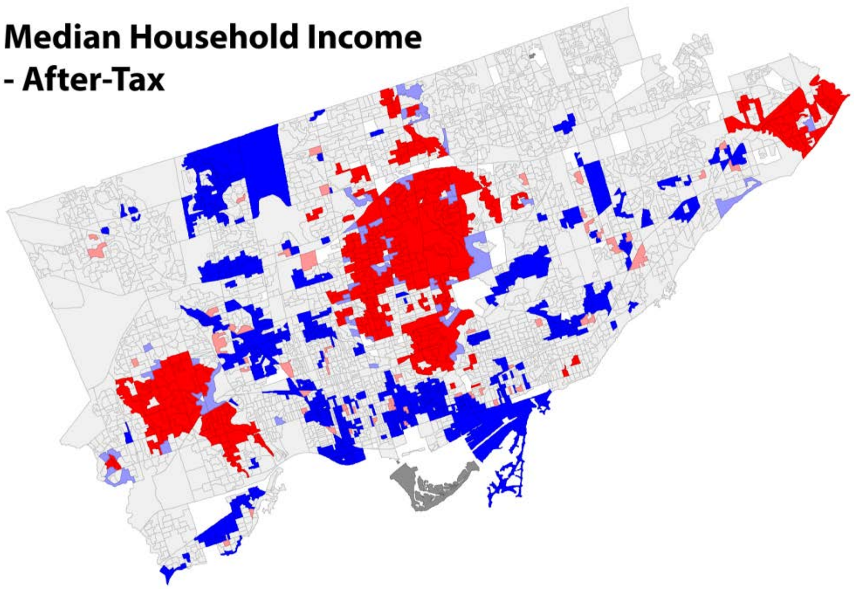
% White Collar



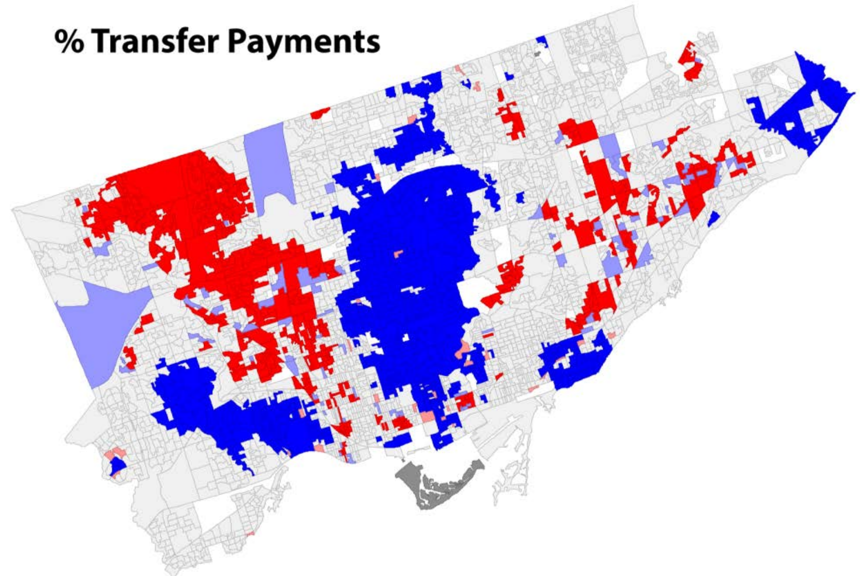
Median Household Income - Before-Tax



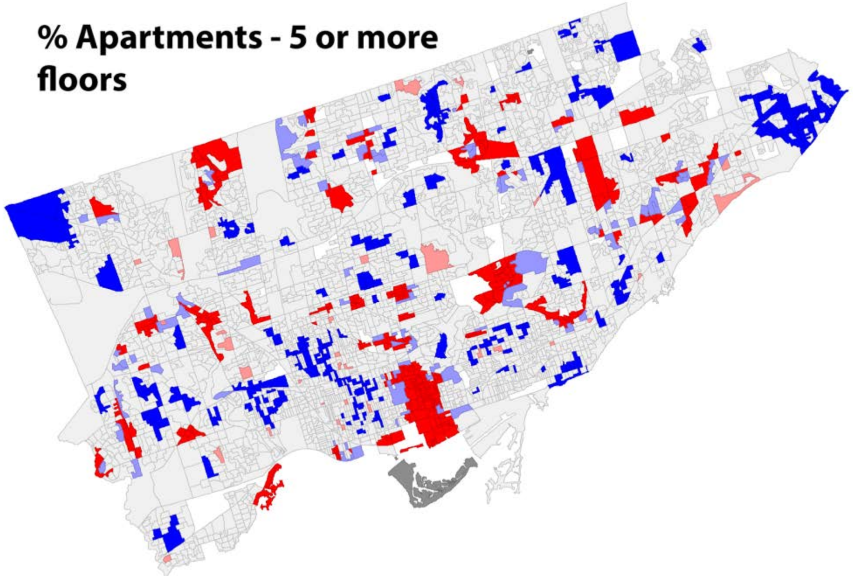
Median Household Income - After-Tax



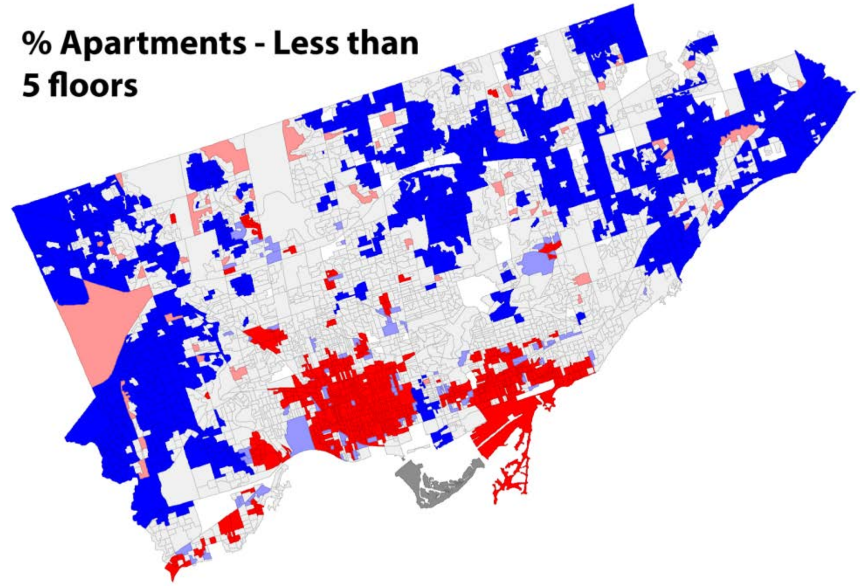
% Transfer Payments



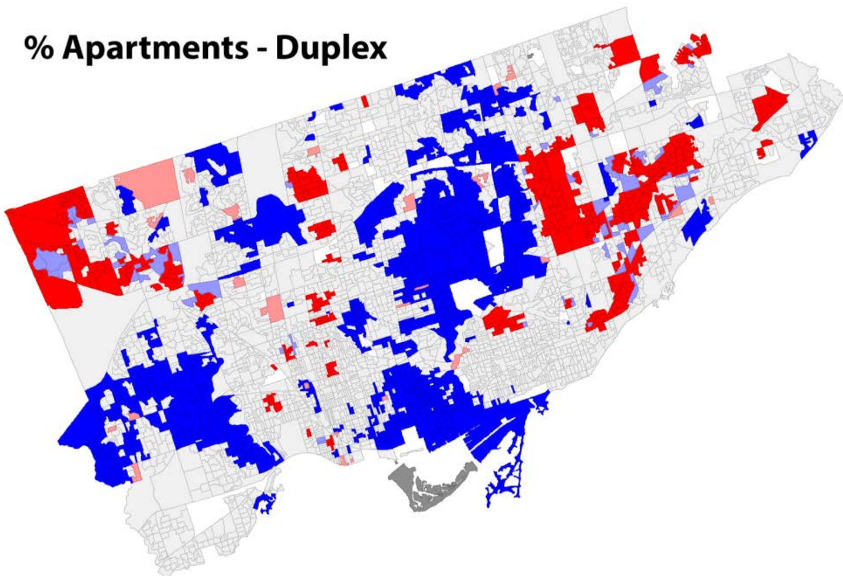
**% Apartments - 5 or more floors**



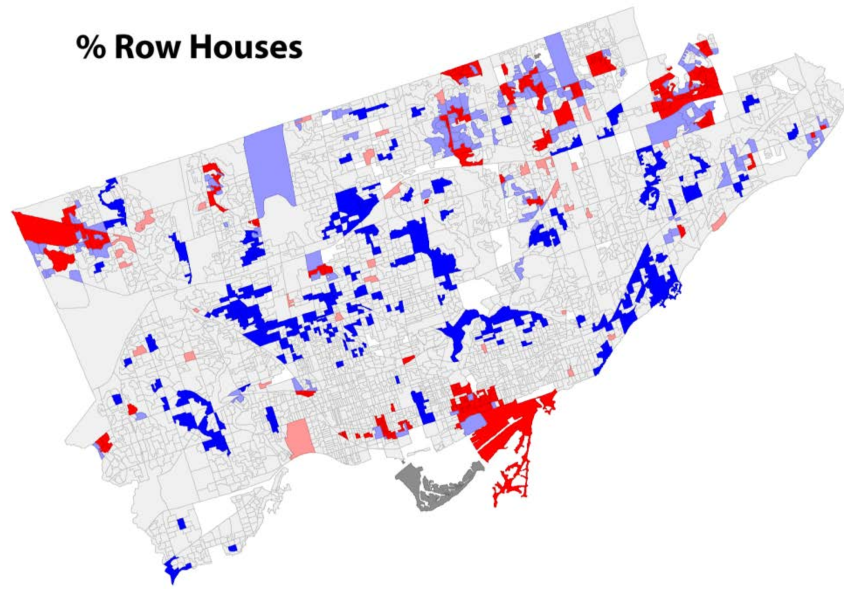
**% Apartments - Less than 5 floors**



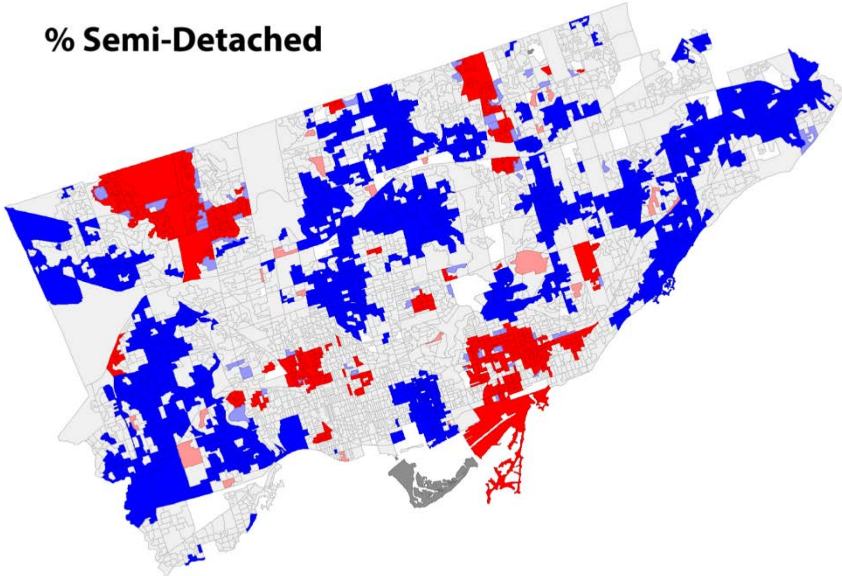
**% Apartments - Duplex**



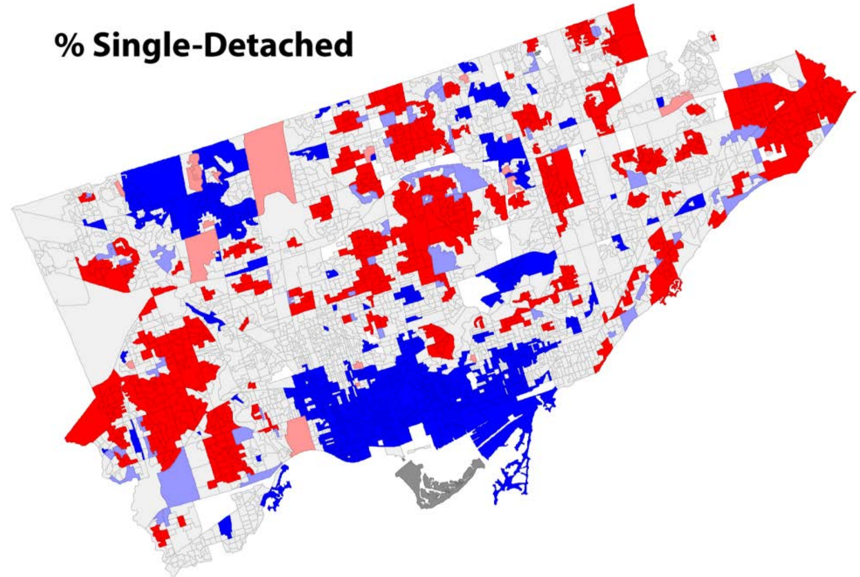
**% Row Houses**



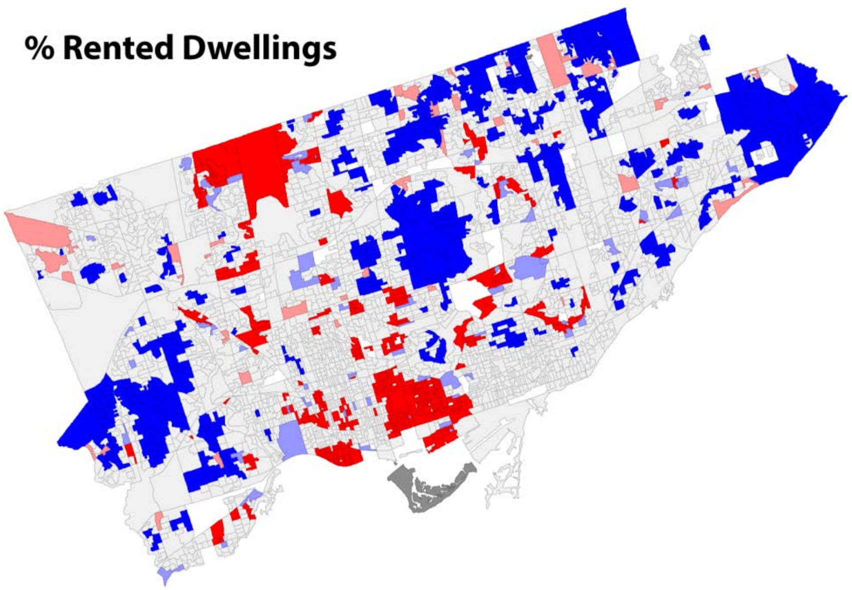
**% Semi-Detached**



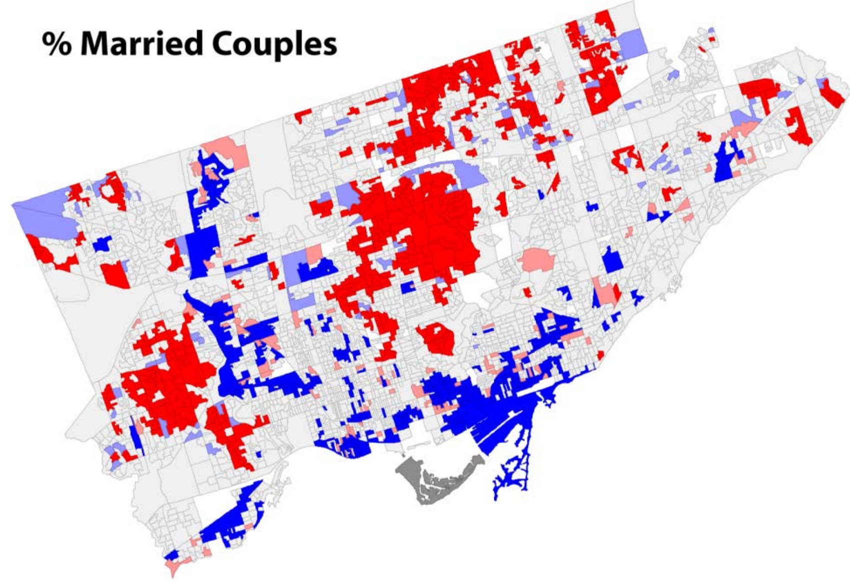
**% Single-Detached**



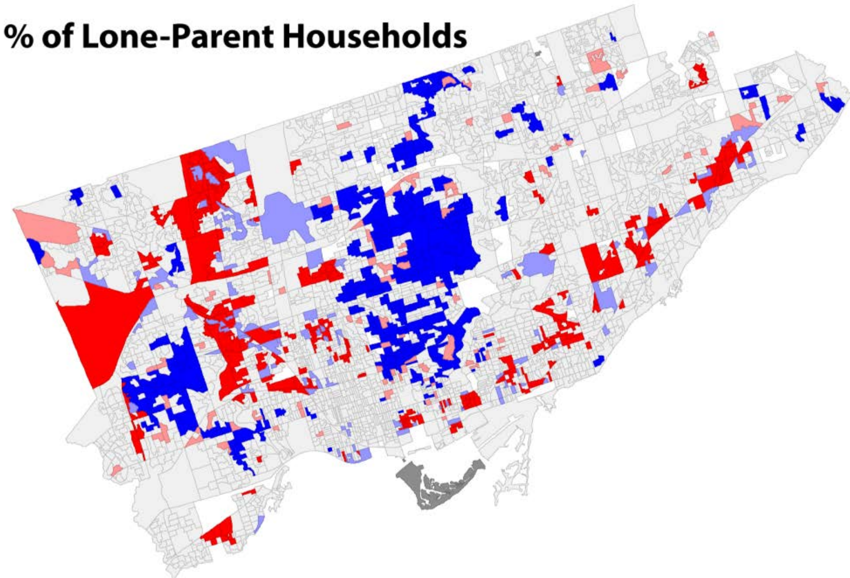
**% Rented Dwellings**



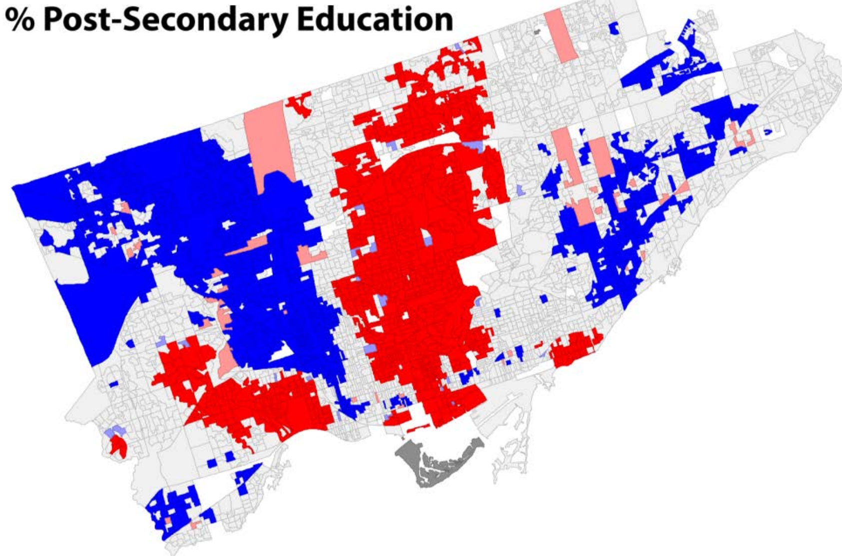
**% Married Couples**



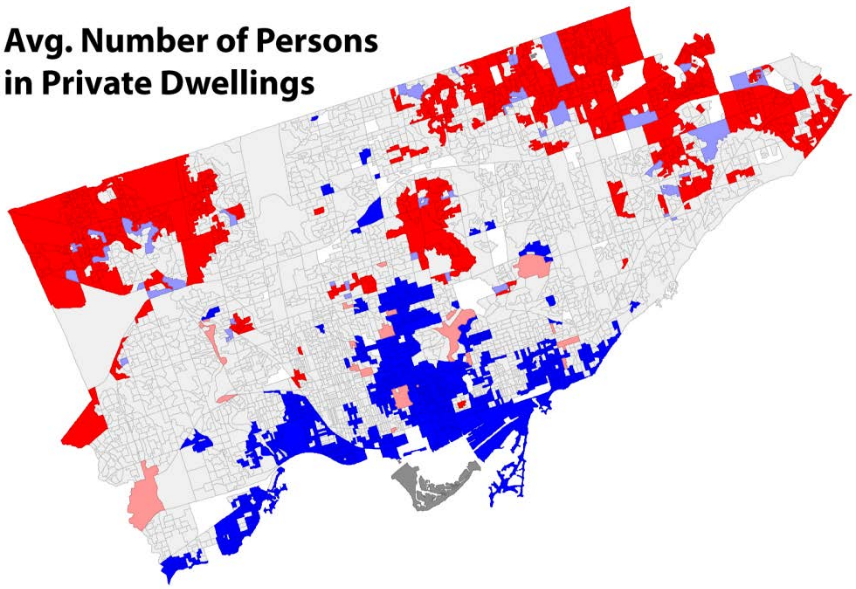
**% of Lone-Parent Households**



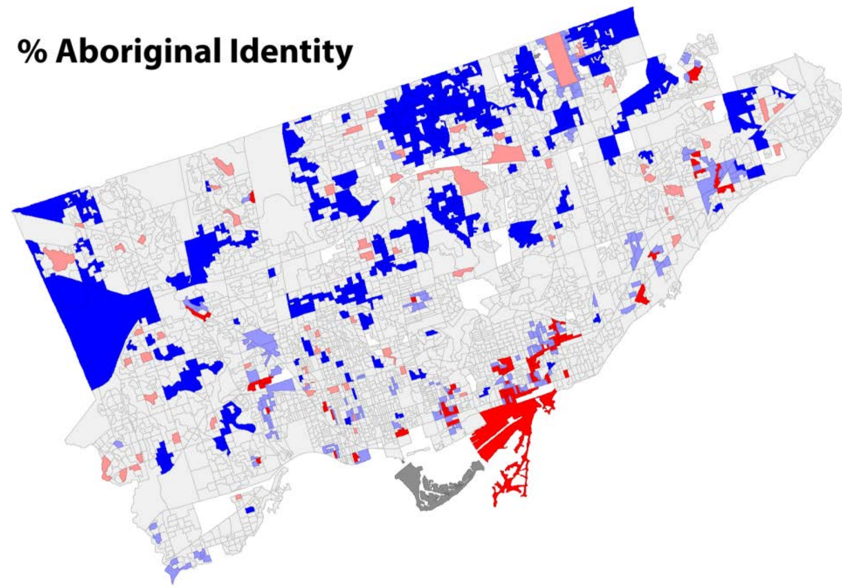
**% Post-Secondary Education**

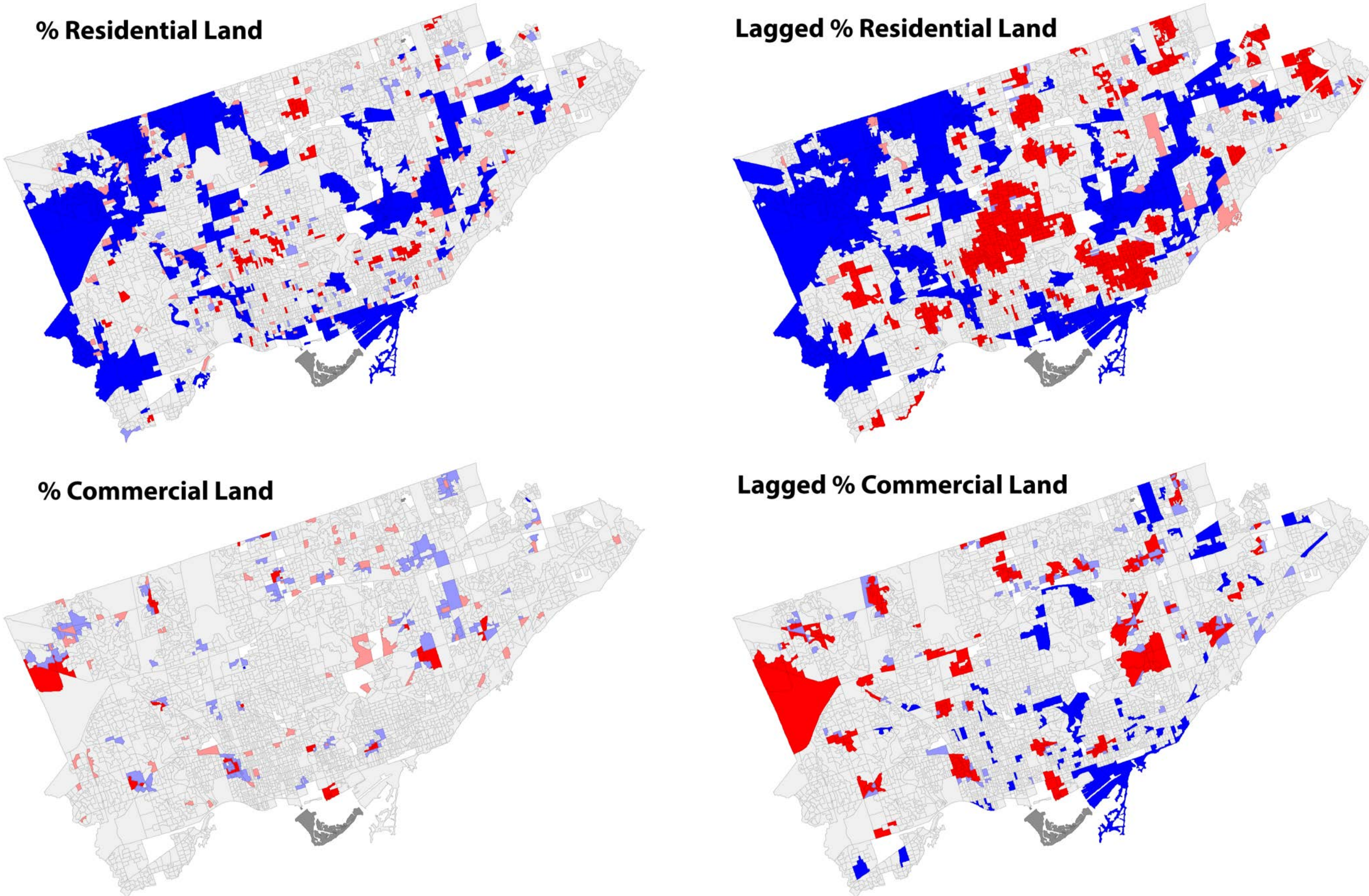


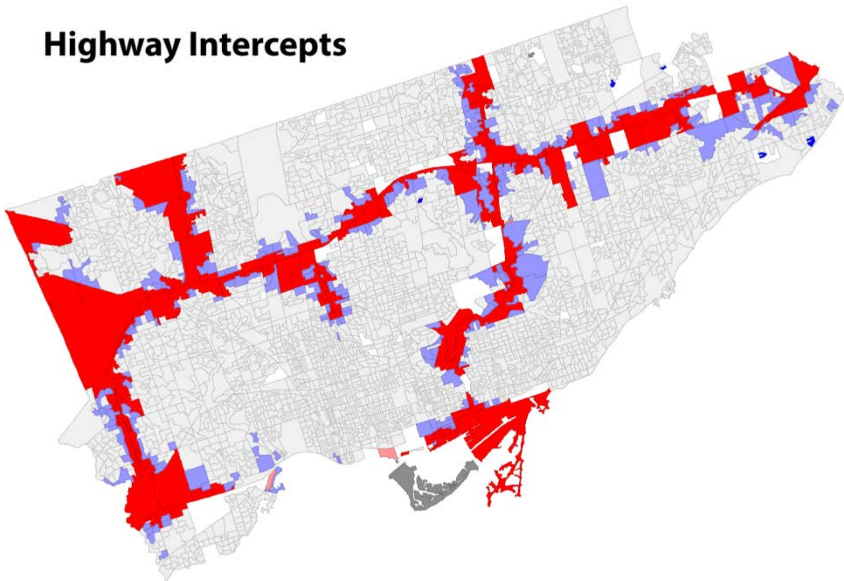
**Avg. Number of Persons in Private Dwellings**

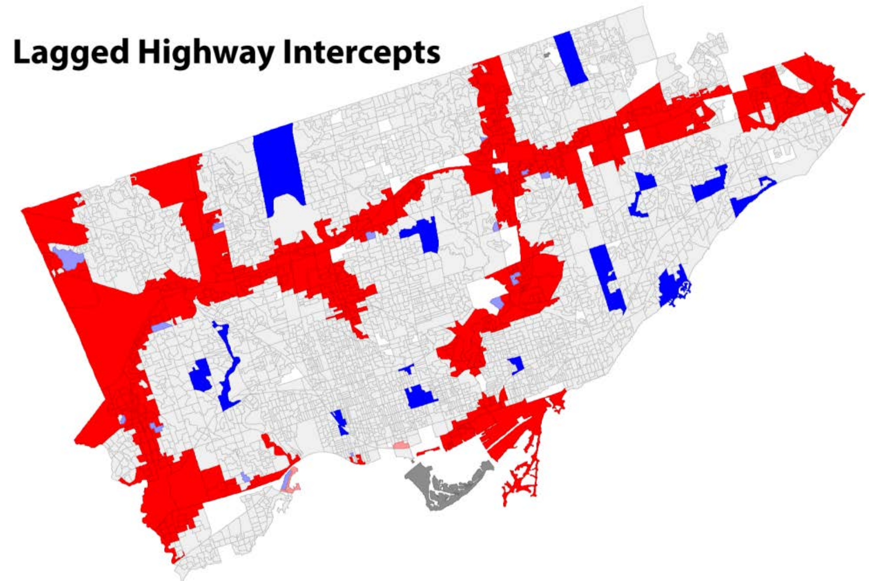


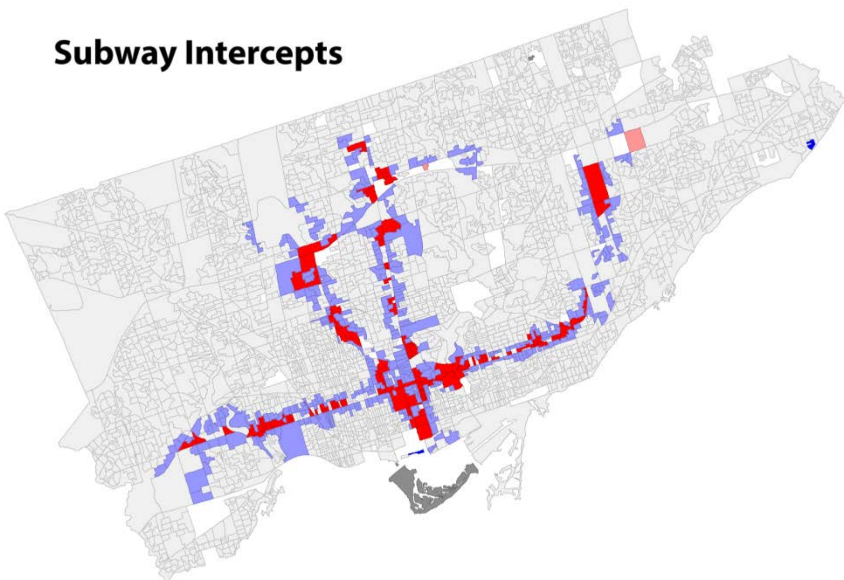
**% Aboriginal Identity**

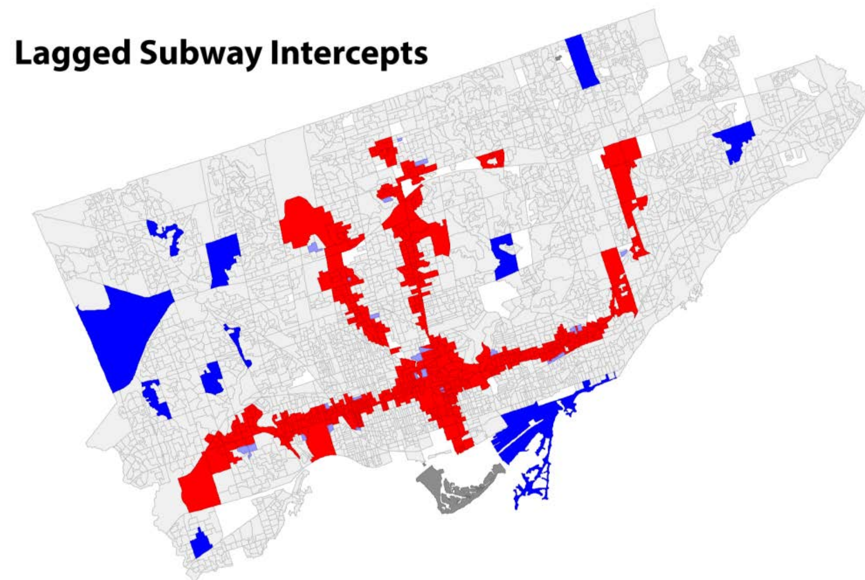


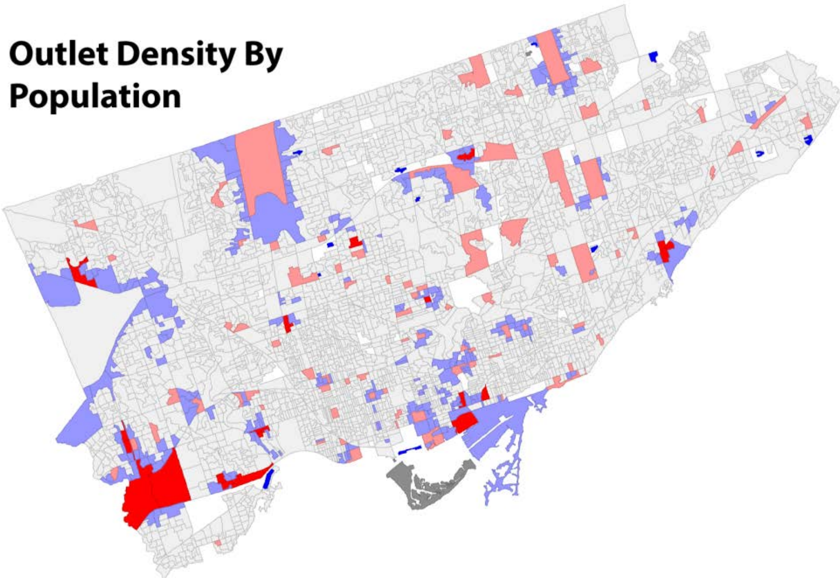

% Residential Land
Lagged % Residential Land
% Commercial Land
Lagged % Commercial Land

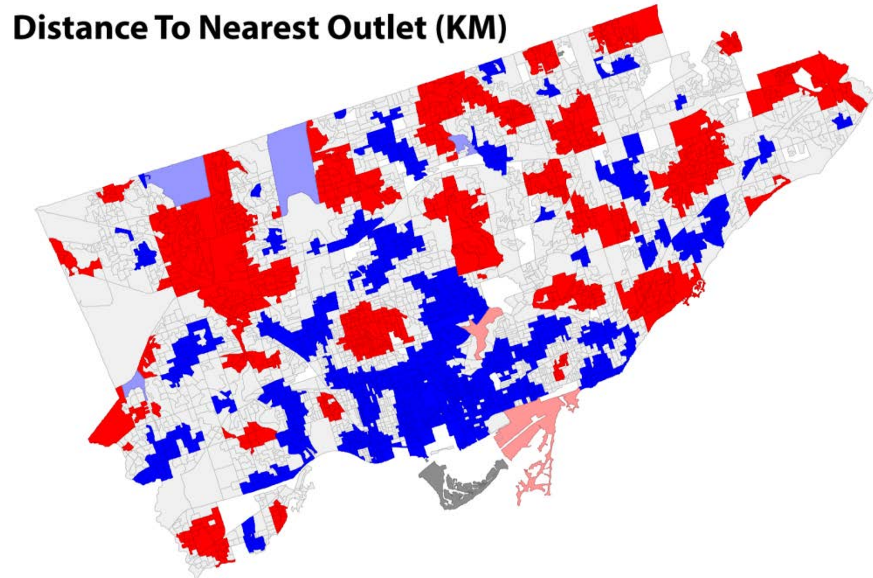

Highway Intercepts

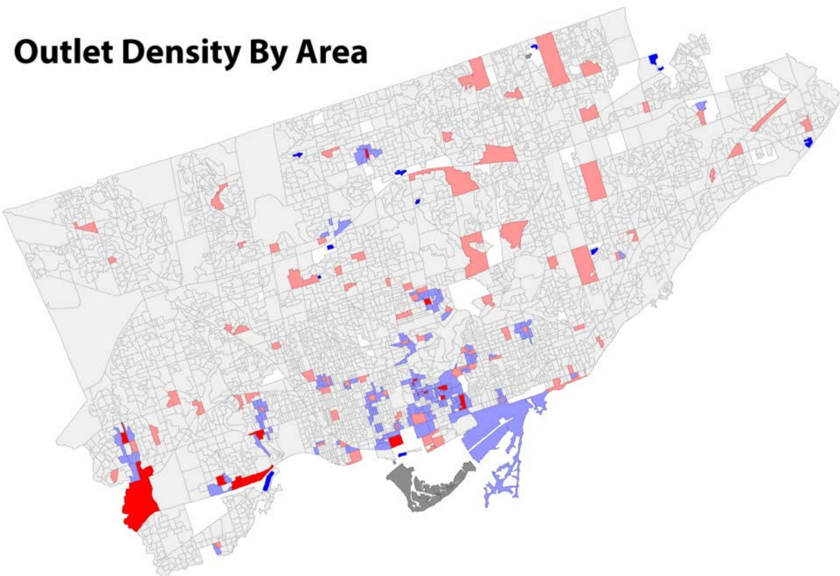

Lagged Highway Intercepts

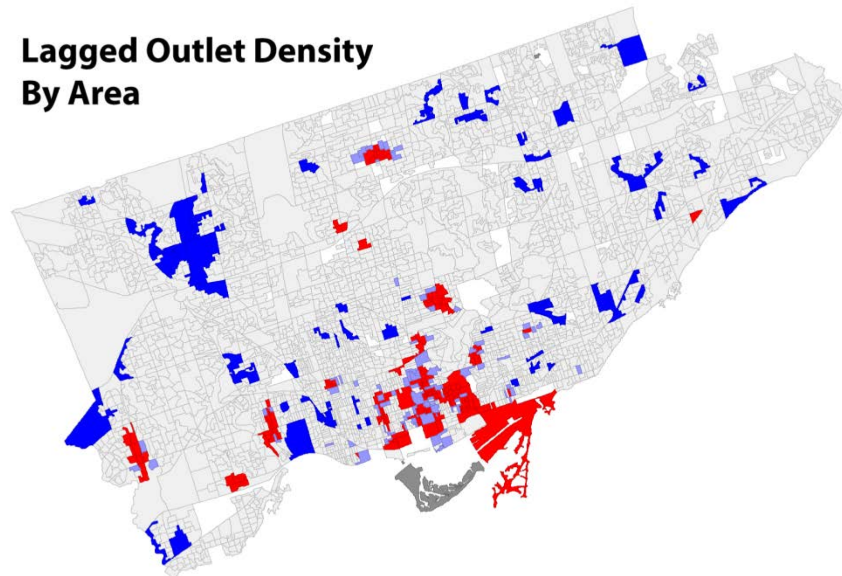

Subway Intercepts


Lagged Subway Intercepts

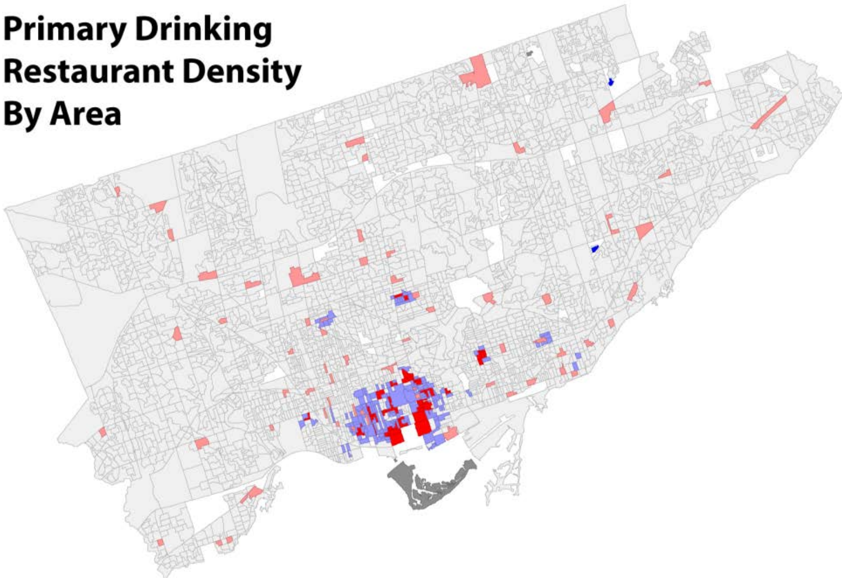

Outlet Density By
Population


Distance To Nearest Outlet (KM)

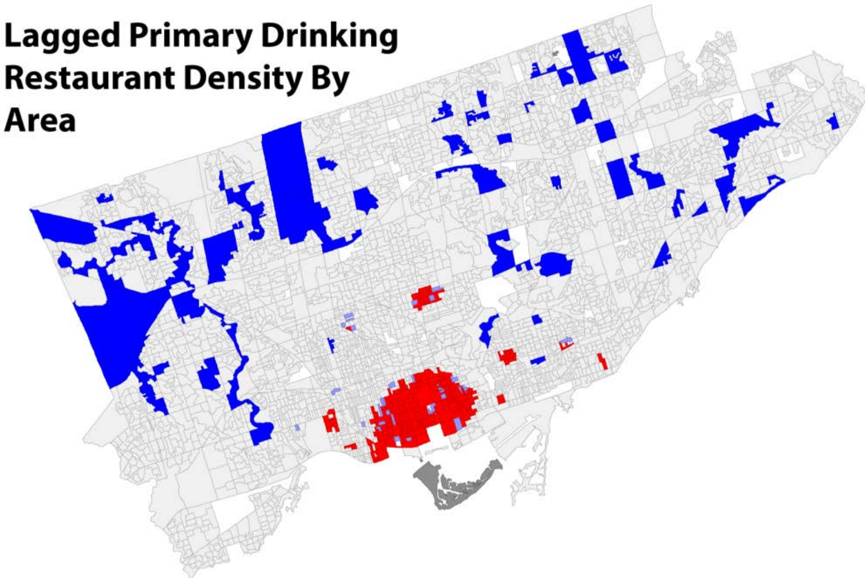

Outlet Density By Area


Lagged Outlet Density
By Area

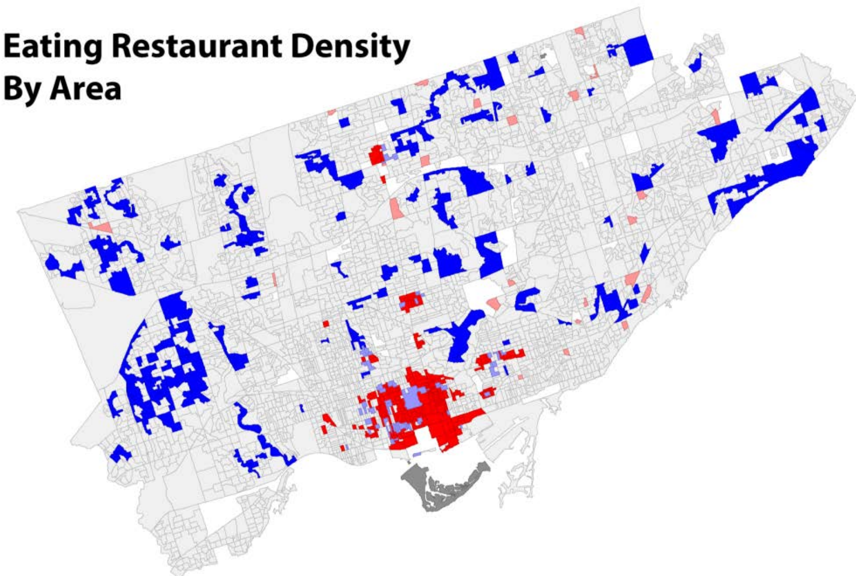
**Primary Drinking Restaurant Density By Area**



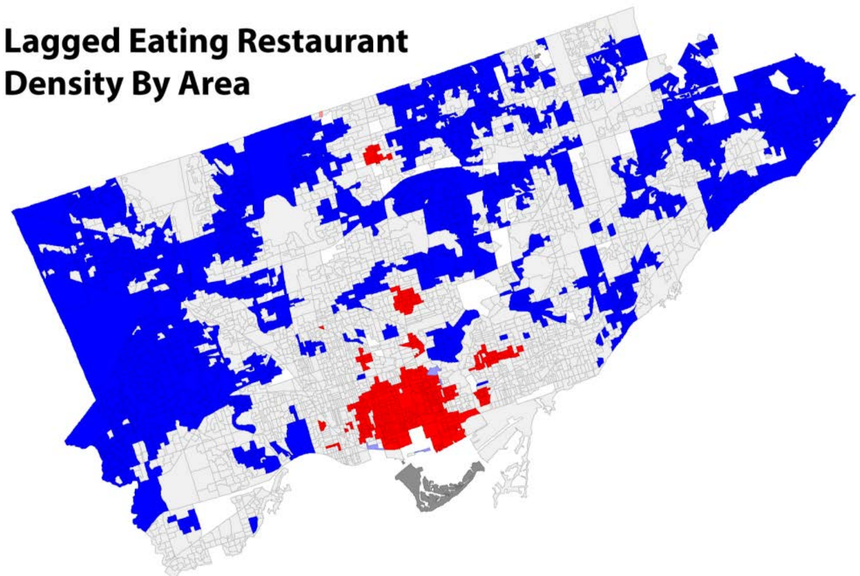
**Lagged Primary Drinking Restaurant Density By Area**



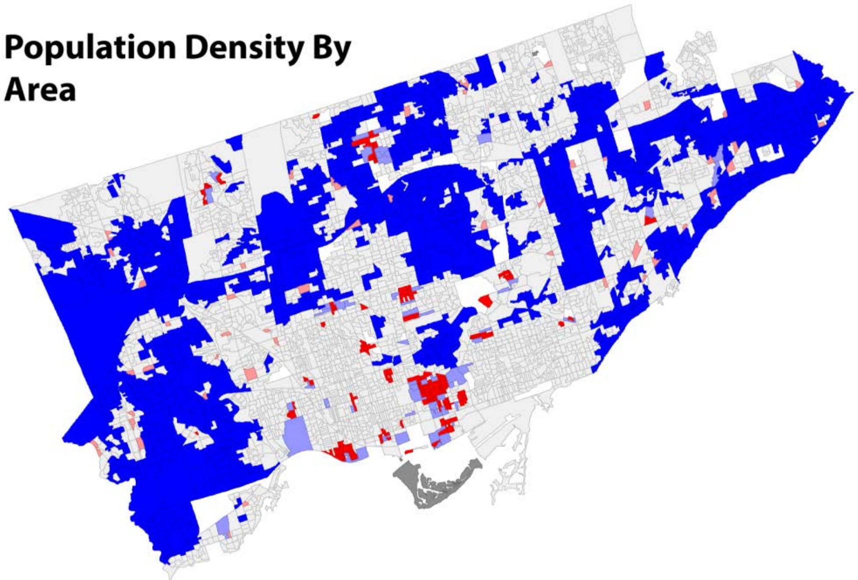
**Eating Restaurant Density By Area**



**Lagged Eating Restaurant Density By Area**



**Population Density By Area**



**Lagged Population Density By Area**