Spatial Analysis of Participation in the Waterloo Residential Energy Efficiency Project

by

Ge Bella Song

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Environmental Studies
in
Geography

Waterloo, Ontario, Canada, 2008

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

Researchers are in broad agreement that energy-conserving actions produce economic as well as energy savings. Household energy rating systems (HERS) have been established in many countries to inform households of their house's current energy performance and to help reduce their energy consumption and greenhouse gas emissions. In Canada, the national EnerGuide for Houses (EGH) program is delivered by many local delivery agents, including non-profit green community organizations. Waterloo Region Green Solutions is the local non-profit that offers the EGH residential energy evaluation service to local households. The purpose of this thesis is to explore the determinants of household's participation in the residential energy efficiency program (REEP) in Waterloo Region, to explain the relationship between the explanatory variables and REEP participation, and to propose ways to improve this kind of program.

A spatial (trend) analysis was conducted within a geographic information system (GIS) to determine the spatial patterns of the REEP participation in Waterloo Region from 1999 to 2006. The impact of sources of information on participation and relationships between participation rates and explanatory variables were identified. GIS proved successful in presenting a visual interpretation of spatial patterns of the REEP participation. In general, the participating households tend to be clustered in urban areas and scattered in rural areas. Different sources of information played significant roles in reaching participants in different years. Moreover, there was a relationship between each explanatory variable and the REEP participation rates.

Statistical analysis was applied to obtain a quantitative assessment of relationships between hypothesized explanatory variables and participation in the REEP. The Poisson regression model was used to determine the relationship between hypothesized explanatory variables and REEP participation at the CDA level. The results show that all of the independent variables have a

statistically significant positive relationship with REEP participation. These variables include level of education, average household income, employment rate, home ownership, population aged 65 and over, age of home, and number of eligible dwellings. The logistic regression model was used to assess the ability of the hypothesized explanatory variables to predict whether or not households would participate in a second follow-up evaluation after completing upgrades to their home. The results show all the explanatory variables have significant relationships with the dependent variable. The increased rating score, average household income, aged population, and age of home are positively related to the dependent variable. While the dwelling size and education has negative relationships with the dependent variable.

In general, the contribution of this work provides a practical understanding of how the energy efficiency program operates, and insight into the type of variables that may be successful in bringing about changes in performance in the energy efficiency project in Waterloo Region. Secondly, with the completion of this research, future residential energy efficiency programs can use the information from this research and emulate or expand upon the efforts and lessons learned from the Residential Energy Efficiency Project in Waterloo Region case study. Thirdly, this research also contributes to practical experience on how to integrate different datasets using GIS.

Acknowledgements

I would like to express my gratitude to a number of people who have provided assistance and support throughout the development and completion of this thesis.

First and foremost, I would like to thank my advisor, Dr. Douglas Dudycha, for his patience and encouragement during the writing process and for his many ideas related to my research analysis. Especially, Doug has taken quite a long time in correcting my wording problems. I appreciate his direction concerning logic clarification and presentation improvement in the thesis greatly. Without his support, this thesis would not have been possible.

I am grateful to my committee member, Professor Paul Parker, for helping me to develop this research topic and giving me a lot of helpful suggestions, helping me get the original data from the REEP, and spending a lot of time on reading my thesis.

Alex Brenning merits special thanks for his help and kindness when I needed advice on statistical analysis.

Finally, but certainly not least, I would like to thank my mother and father for their undying support and encouragement during these past few years. This thesis could not have been completed without their understanding.

I would like to thank my best friend, Hao Zhang who was always willing to listen to the unending rants and complaining that comes with an anxious and nervous graduate student. Thanks also go out to my good friend Holly Waite who always helped edit my writing and correct the grammar.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Problem Statement

Climate change is one of today's most serious sustainable development problems. The increased concentration of greenhouse gases in the earth's atmosphere is a cause of global warming and also leads to other serious problems, such as deterioration of ecological systems worldwide. The residential sector in Canada accounts for approximately 17-20% of national greenhouse gas emissions (Harchaoui, 2003; Parker, et al., 2004). Understanding the impact of energy efficiency on the environment has become increasingly important in recent years, as countries and regional communities realized energy efficiency needed to be improved by encouraging households to minimize human-caused climate change and at the same time have healthier homes and sustainable communities (Parker, et al., 2004). Through an energy efficiency program, the residents who participate in the program can reduce energy use and energy cost, help improve their local air quality, and contribute to the global effort to reduce climate change. In Canada, the federal/provincial National Action Program on Climate Change (NAPPC) calls for cooperation and action by all levels of government to implement energy efficiency programs to reduce greenhouse gas emissions.

This research focuses on exploring the factors that influence the participation of households in the Residential Energy Efficiency Project (REEP) in the Regional Municipality of Waterloo hereinafter called Waterloo Region. In order to identify potential test factors, this thesis reviewed the existing body of research that links the topics in public behaviors toward energy conservation and barriers to improving energy efficiency. Test factors were selected based on the literature review and data availability to analyze which factors influence participation in energy efficiency programs. Geographic Information Systems (GIS) analysis, Poisson regression and logistic regression were employed in this research to analyze participation in the Residential Energy Efficiency Project.

1.2 Background information for Energy Efficiency Programs

What is energy efficiency?

The World Energy Council (WEC) defined Energy Efficiency as "encompasses all changes that result in a reduction in the energy used for a given energy service (heating, lighting...) or level of activity. This reduction in the energy consumption is not necessarily associated to technical changes, since it can also result from a better organization and management or improved economic efficiency in the sector (e.g. overall gains of productivity)"(World Energy Council, 2008).

More simply, Energy efficiency means using less energy while having the same productivity or doing the same job. For example, doing the laundry, lighting and heating the home or office, and processing wood products or building a car all use energy. In most situations, all these things can be done by using much less energy, with better insulation in buildings, more efficient appliances and lighting and process improvement to save money and energy.

The benefits from EEP

Implementing an energy efficiency program (EEP) for a local community can bring many benefits, such as creating jobs when the EEP start to operate; reducing the impact to the environment because less energy use can save more greenhouse gas emission; saving money for homeowner, businesses, industry and the municipality from the energy bill; reducing production costs for industry and businesses; reducing exposure to rising energy prices and energy shortages; winning new customers when someone considers whether the products are environment friendly or not. (World Energy Council, 2008; City of Toronto, 2008; Residential Energy Efficiency Project, 2007b)

Residential Energy Efficiency Project (REEP)

"The Residential Energy Efficiency Project (REEP) provides the citizens of Waterloo Region with tools for sustainable resource use, energy conservation and efficiency, and renewable energy applications" (REEP, 2007b).

REEP is an environmental non-profit program created by a partnership between the Faculty of Environment at the University of Waterloo and the Elora Centre for Environmental Excellence. The mission for REEP is working towards healthier homes and sustainable communities (REEP, 2007b). The program is currently offered by Waterloo Region Green Solutions, the delivery agent in Waterloo Region (Ontario, Canada) for the national *EnerGuide for Houses* (EGH) program. EGH was designed and developed by the federal government and administered by the Office of Energy Efficiency (OEE), in Natural Resources Canada (NRCan) (Natural Resources Canada, 2007a). Waterloo Region Green Solutions is a member of Green Communities Canada, a national association of non-profit environmental service organizations that share programs and resources including the contract to deliver EGH in Ontario.

REEP started operations in May, 1999 and stopped delivering initial EGH evaluations in April 2006 when the federal government withdrew funding and cancelled the program. However, follow-up evaluations to document the improvements achieved by households continued during 2006 and in 2007 the federal government introduced a similar program under a new name, ecoENERGY. Waterloo Region Green Solutions continues to provide residential energy evaluations under the REEP banner to maintain consistent messaging in the community.

The REEP service provides an evaluation of the energy efficiency of houses and identifies steps to improve their efficiency. Once the home owner books an appointment for a home energy efficiency evaluation, a Certified Energy Advisor will come to that home and use a computer modeling system to determine the home's energy efficiency performance and identify sources of heat loss. Generally, the evaluation takes 2 to 4 hours to complete, depending on the size of the home (REEP, 2007a). After the evaluation is complete, the Advisor will provide a detailed report to the homeowner containing the recommended parts in the home that need upgrading, energy savings estimate after upgrades, and an *EnerGuide for Houses* label and rating from 0 to100 that represents the level of efficiency of the home (see Appendix 6). Also, if participants would like to be eligible for government grants, they have to do a second evaluation (follow up evaluation) within 18 months and submit it to NRCan. Sometimes participants may take longer than this time to do the retrofits of their homes. The information from follow up visits can be used to assess the results of energy efficiency upgrades. "Natural Resources Canada surveys show that most customers implement some energy saving measures after an EnerGuide for Houses visit, even if they don't have the results assessed through a follow-up visit" (REEP 2006 Annual Report, 2007).

The benefits from REEP include: reduced greenhouse gas emission; improved local air quality by using less energy for homes; local economic stimulus through upgrades of evaluated homes (like labor and materials); job opportunities; and information for research on home energy use (Kitchener Utilities, 2005). The eligible homes for REEP are low-rise residential properties, either single detached, duplexes, townhouses or row houses. Multi-storey apartments are not eligible (Kitchener Utilities, 2005).

1.3 Research Goals and Objectives

The primary goals of this thesis are:

1. To understand household behavior toward energy conservation;
2. To explore the spatial distribution of participating households each year and overall and assess the marketing methods that play an important role in influencing participation;
3. To determine the factors affecting participation in the residential energy efficiency program in the Waterloo Region;
4. To make recommendations to encourage greater participation in future programs of this type, and for future research.

Four research tasks are carried out in order to achieve the above stated goals:

1. Review the existing literature about public attitudes toward energy savings, factors that may influence energy consumption behavior, and barriers to participation in the residential energy efficiency program;
2. Undertake spatial analysis and visual comparison by using Geographic Information Systems (GIS) to create thematic maps based on the data available from Statistic Canada and REEP office, posing four research questions:
    a) What is the participation rate based on CDAs level;
    b) What is the spatial pattern of participating households;
    c) Is the distribution of participation households affected by sources of information like referral, insert, or direct market, and so on;
    d) How did the spatial pattern of participation change over time (from 1999 to 2006);
3. Based on the literature review, identify explanatory factors that may explain the participation rate, and create a series of thematic maps in GIS;
4. Undertake statistical analysis of quantitative data generated by the REEP, Statistics Canada and ArcGIS to determine whether relationships exist between hypothesized explanatory variables and the number of participating households or the participation rate.

1.4 Thesis Structure

The thesis consists of six chapters. Chapter two reviews the current status of energy efficiency programs. Chapter three discusses the methodological approach taken to do the visual comparison and spatial analysis, including a description of the study area, main objectives and initial hypothesis, data processing required to build the GIS database used for the analysis, results of visual interpretation of the data and discussion of the findings. The chapter four provides first statistical method, Poisson Regression, to explore the form of the relationship between the hypothesized explanatory variables and REEP participation. The chapter five discusses the second statistical method, Binary Logistic Regression. This method was used to analyze what factors affect whether or not the participating households take the second (follow-up) evaluation of their homes. The final chapter synthesizes findings from a review of the energy efficiency literature and results based on the three analysis methods regarding the spatial pattern of participation in the REEP and relationships between participation and hypothesized explanatory variables. This chapter also summarizes the contributions of this study to the literature and outlines recommendations for further research.

Chapter 2 Literature Review

The following literature review is intended to enhance the understanding of the theories and practice that underlie the purpose and objectives of this study. This chapter will focus on a review and study of the residential sector in the energy efficiency initiatives. It is essential to understand public attitudes toward energy efficiency programs, and to identify the key barriers which may govern participation for energy efficiency programs. Based on that knowledge, energy efficiency programs can be adjusted to gain more participants. Through their participation they can live in a healthy environment and conserve the natural environment by reducing GHG emissions. Understanding residential energy use behavior is necessary for successfully promoting residential energy efficiency programs. The first part will introduce the energy efficiency initiatives taken by different levels in Canada, and is followed by a review of the determinants of energy conservation behaviors, and identifying the barriers to participation in energy efficiency programs and the use of energy efficient products.

2.1 The Federal, Provincial and Community Energy Efficiency Initiatives

Improving energy efficiency reduces GHG emissions which benefits the natural environment, improves human living standards and the economy and also contributes to energy security. However, achieving these benefits requires the involvement of all levels of government to implement successful energy efficiency programs that encourage all citizens to participate and to use energy efficiency products.

2.1.1 Federal Government Initiatives

The federal government of Canada has an office called *the Office of Energy Efficiency (OEE)*. This office has become Canada's primary center for collecting, analyzing and delivering key programs that promote energy efficiency in the major energy using sectors of the economy (Natural Resources Canada, 2006b). The OEE helps Canadians make energy efficient choices when buying, selling or manufacturing energy-using equipment (Natural Resources Canada, 2006b). Since 1990, the NRCan has made annual assessments of trends in energy use and related greenhouse gas (GHG) emissions in Canada. They have published the results in technical reports called *Energy Efficiency Trends in Canada*. The OEE feels the changes in energy efficiency cannot be directly measured at the sectoral or economy-wide level, so they developed the OEE Energy Efficiency Index by using a methodology call factorization. "This index is only an estimate of changes in energy efficiency in the economy" (Natural Resources Canada, 2006b). Natural Resources Canada (2006a) shows that energy efficiency improved by 14 percent from 1990 to 2004, and as a result of this improvement, Canadians saved around $14.5 billion in energy cost in 2004 alone, and reduced greenhouse gas emissions by 53.6 megatonnes.

There also have been some regulations under OEE, for example, EcoENERGY Efficiency. The objective of this regulation is to improve the energy efficiency of products and equipment in Canada by amending the Energy Efficiency Regulations established to implement the goals of the Energy Efficiency Act (Natural Resources Canada, 2006b). Regulations under the Energy Efficiency Act have been in effect since 1995 and set minimum acceptance levels for a number of energy-using products such as appliances and heating, lighting, and air-conditioning products. Energy efficiency action through regulation means the government of Canada can address harmful GHGs and air pollutants through its regulatory powers. On October 19, 2006 the government tabled amendments to the Energy Efficiency Act in Parliament, and it took concrete action to improve the standards for a range of consumer products and equipment (Natural Resources Canada, 2006c). “Broadening and strengthening the Act means that 80 percent of the energy used in homes and businesses will soon be regulated” (Natural Resources Canada, 2006c).

2.1.2 Provincial Energy Efficiency Initiatives

Each province designs their own programs which are suitable for their situation. For example, Ontario has a program called “Energy Efficiency Assistance Program for Houses ON” (OPA, 2008). The main goal of this program is to develop saving of 1,300 megawatts of electricity over the next several years, or enough electricity to supply 500,000homes (Green Communities Canada, 2008; OPA, 2008). One of these initiatives focuses on achieving cost effective savings of 100MV of electricity, or enough power for 33,000 houses (Green Communities Canada, 2008).

The Energy Efficiency Assistance Program for Houses (EEAPH) is intended to help income-eligible households improve the efficiency of their homes and to reduce their electricity bills. Only households in owner occupied and private rental housing where the residents pay the energy bills are eligible for this program, however, social housing is not eligible for this program.

A quick start energy efficiency strategy for Ontario has been proposed by Winfield & Hall (2006). In this report, they outlined the energy efficiency strategy for Ontario. The strategy is expected to be put in place over the next two years. This report extracts the experience in the US with successful state level energy efficiency programs, and outlines three initiatives and nine specific programs to be undertaken in Ontario (Winfield & Hall, 2006). The Quick-Start strategy aims at savings in grid-electricity consumption of 23,000 GWh/yr, with an implied reduction in required generating capacity of 4500MW, by 2012 (Winfield & Hall, 2006).

Some other provinces also have programs similar to Ontario, to make energy conservation and environmental savings, such as Yukon and Alberta. The Yukon government has an agency, Yukon Housing Corporation, that announced six new residential energy programs in 2007 which offer zero percent interest rates on loans for home repair, energy efficiency items, alternative energy systems under the residential energy management program, home energy evaluation, and green home incentive for existing homes to enhance the energy efficiency performance of Yukon's housing stock (Yukon Housing Corporation, 2007). Alberta Furnace Replacement Program is an incentive program that offers rebates for consumers to upgrade their furnace to a high efficiency furnace. The Alberta Washing Machine Program offers rebates as well for consumers to upgrade their washing machines to an Energy Star approved one (Natural Resources Canada, 2007b).

2.1.3 Community Energy Efficiency Initiatives and the Benefits

Sometimes, cities feel they have to wait too long for federal or provincial action and therefore choose to promote their own energy efficiency programs. An example is EnviroCentre that delivers a service similar to EEAPH in the Ottawa area (refer to section 2.1.2). EnviroCentre is a non-profit partner for delivering energy-efficiency services in the City of Ottawa (EnviroCentre, 2007). EnviroCentre works through a public-private partnership with the City to deliver energy efficiency goods and services (EnviroCentre, 2007). This Organization is best known for its work on residential energy efficiency, to deliver the federal Energuide for Houses program showing its retail clients how to save energy and money while making their homes more comfortable and live in a healthier environment.

The City of Vancouver also has incentive programs for energy efficiency very similar to the REEP. The government rebates offer monetary for home energy improvements. Vancity provides Bright Ideas Financing for home improvements that incorporate energy efficiency through the ecoENERGY program. If the home' increases a minimum of five points on the EnerGuide rating, then this home may be "eligible for a low-interest renovation loan (at prime rate), for up to ten years and in the range of $3,500-$20,000" (City of Vancouver, 2008).

The Energy Efficiency Office (EEO) is responsible for developing an energy efficiency and conservation strategy for Toronto. This has helped to reduce carbon dioxide emissions by 20%, relative to 1988 levels in the City of Toronto. The initiatives for the EEO to achieve this goal include the Better Buildings Partnership-Existing Buildings, the Better Buildings Partnership-New Construction and creating the Energy Plan (City of Toronto, 2008). The Better Buildings Partnership - Existing Buildings (BBP-EB) assists building owners improving their buildings through energy efficiency measures. "The Better Buildings Partnership New Construction Program (BBP-NC) goal is to have new buildings designed and built to be more energy efficient than those designed to only meet the minimum requirements of the Ontario Building Code"(City of Toronto, 2008). A few of the benefits to build energy efficient buildings are include, lower utility bills, reduce greenhouse gas emissions, increased resale value, less strain on our energy supply (City of Toronto, 2008). In July 2007, the Climate Change, Clean Air and Sustainable Energy Plan approved by City Council to help the city improve their air quality and slow down climate change caused by coal and fossil fuels, lower the energy costs, etc (City of Toronto, 2008).

The energy efficiency of buildings can be improved through the use of certain materials such as attic insulation, components such as insulated windows, and design aspects such as passive solar or passive house features using renewable energy sources (REEP, 2007a). Further, the energy efficiency of communities and cities can be improved through architectural design, and transportation system design. Thus, energy efficiency involves all aspects of energy production, distribution, and end-use.

Now, concerned about waste energy, many communities are engaged in residential energy conservation programs for existing houses that are poorly designed (Cohen, Goldman, & Harris 1991; Schweitzer, Hirst, & Hill, 1991; Brown, Berry, Bdzer, & Faby, 1993; Clark & Berry, 1995). Energy efficiency programs offer more than energy benefits (Tonn & Peretz, 2007). Energy efficiency programs can benefit all households and especially low income households. Other benefits from energy efficiency programs include environmental benefits, community benefits, and benefits to landlords (Stewart & Fry, 2006; Tonn & Peretz, 2007; Residential Energy Efficiency Project, 2007a; Energy Savings Plan, 2007).

Lower energy bills means that housing is more affordable which can benefit low income people, because they are seeking a rental fee for accommodation cost lower than 30% of their income, so reduced energy costs through energy efficiency program or products also can help to improve shelter affordability (Stewart & Fry, 2006). Other benefits to low income households like more comfortable housing, reduced chance of homelessness and stress over energy bills, reduced risk of accidents from supplementary heating can be achieved as well (Stewart & Fry, 2006; Tonn & Peretz, 2007). Tonn & Peretz (2007) and Schweitzer & Tonn (2001) found that energy efficiency program without these non- energy benefits also add benefits to local economy, for example, increased local employment through an energy efficiency program and increased property values as well. Usually, energy efficient homes increase resale value by 2% to 8% compared to less efficient homes (Residential Energy Efficiency Project, 2007a; Energy Savings Plan, 2007).

Environmental benefits include reduced pollution, due to conservation and increased efficiency using the cleanest way to meet our energy needs, and less reliance on coal, gas, nuclear energy and other nonrenewable sources (Stewart & Fry, 2006).The benefits also can contribute to community, for example, energy efficiency can reduce costs to shelters, social service agencies and the health care system due to avoided homelessness and more affordable housing, especially for low income people (Stewart & Fry, 2006). Money not being spent on energy waste can be re-invested in the local community.

The Stewart & Fry (2006) study indicates there are benefits to landlords as well. Landlords who make investments to improve the energy efficiency will increase the overall value for their building, reduce turnover of tenants because they would like to stay in the building with low utility costs, and the rental fees are more affordable that means the tenants can pay on time since less money needs to be paid for energy costs and it becomes easier to meet the rent.

2.2 Determinants of Energy Conservation Behaviors

Understanding what variables may influence energy conservation behaviors is very important to allow for improving energy conservation and developing energy efficiency programs. Previous studies have investigated many different variables to determine their influence on energy consumption. These include demographic, economic, and structural variables, as well as personal variables such as attitudes and beliefs. (Emery & Gartland, 1996; Gatersleben, Steg, & Vlek, 2002; Guerin, Yust, & Coopet, 2000; Seligman, Darley, & Becker, 1978; Kempton, et al., 1985; McMakin, Malone, & Lundgren, 2002; Wilhite & Ling, 1995).

Guerin et al. (2000) mentioned a human ecosystem model developed by Guerin (1992) and adopted from Bubolz, Eicher, & Sontag (1979).The human ecosystem model was developed to examine the relationship of the human, and the natural, social, and designed environments in the context of energy consumption (Guerin et al., 2000). There are four parts in this model, three environments and the human as a separate part of the system. However, the human can affect the environment or the environment can affect the human (Guerin et al., 2000). Guerin et al. (2000) and Yust, Guerin, & Coopet (2002) explain the environments as follows: the natural environment is the physical and biological components; the human represent an individual (occupant), family, or household; the social environment is the psychological and social behaviors of the occupants; the designed environment is anything that is constructed or built by humans. The human ecosystem model's method of categorization will be used for the remainder of this section as a framework to communicate the findings of other energy conservation studies. This section only discusses the human, social environment, and designed environment.

2.2.1 Household Characteristics

A large amount of research has been done to study the relationship between residents' age and their likelihood to conserve energy. Households that participate in an energy audit tended to be younger than nonparticipants (Hirst & Goeltz, 1984). Peter (1990) found that if the households have children under 6 or have elderly members, they were not likely to change their temperature setting. Weihl & Gladhart (1990) have similar findings that if the households have infants and elderly people, usually they were more likely to keep higher temperatures in order to maintain health. Wilhite & Ling (1995) found that younger respondents in their study were more likely to reduce energy consumption than elderly respondents. Warriner (1981) has found that people under 65 used around one third more electricity on the average than people older 65 years old. However, during the winter, the older people used more electricity than younger adults on the average. This may indicate that older adults have a lower tolerance for cold, which results in additional heating requirements (Guerin et al., 2000; Yust, Guerin, & Coopet, 2002).

Black, Stern, & Elworth (1985), and Guerin et al. (2000) found that home ownership is a significant predictor: owner occupants are more likely to engage in energy conservation behavior. Homes occupied by owners are more likely to have energy conservation features because of the personal benefits from the investment, and they are more responsive to long term capital investment, since payback may require a relatively long time (Black, Stern, & Elworth, 1985; Kasulis, Huettner, & Dikeman, 1981; Tienda & Aborampah, 1981)

Households who participate in an energy audit tend to have more education than nonparticipants (Berry et al., 1981). Black et al. (1985) indicate that people with higher education had a better understanding of energy conservation, Johnson-Carroll et al. (1987) also found that people with more education may be more concerned about the seriousness of the energy situation. Gatersleben et al. (2002) analyzed data from a survey among 1250 households on consumer behavior and quality of life. They found that level of education was not significantly related to household energy use. However, Wilhite & Ling (1995) found that respondents with a higher level of education (defined high education as having 12 years of education or more) were more likely to reduce energy consumption, although statistically, this association was not as strong as the association with age. Even though some research shows that education level was not significantly related to household energy use, there were more researchers who found that education level was significantly related to household energy use. The research provides reasonable evidence to support the hypothesis that education level is significantly related to household energy use. Well-educated consumers tend to have higher income and more energy consumptive dwellings, while less-educated consumers tend to have lower incomes and cannot afford high energy consumptive behavior.

One of the other variables thought to influence energy conservation is total household income (Gatersleben et al., 2002; Brandon and Lewis, 1999). Households that participate in an energy audit tend to have higher incomes than non-participants (Berry et al., 1981). Many researchers also identified that the higher the income the greater the potential for energy conservation, especially to conserve electricity (National Research Council, 2005; Nielsen, 1993). Higher income households are more likely to improve equipment efficiency at their homes; since they can afford the larger financial investments that are required to purchase energy saving products and expensive energy efficiency technologies (Straughan and Roberts, 1999; Guerin et al., 2000). Gatersleben et al. (2002) did two case studies, in which they found that energy use appeared to be most strongly related to household characteristics, especially income. They stated that respondents with a higher income used more energy than low income. On the other hand, Johnson-Garroll et al. (1987) found great variations in home ownership, lifestyle and energy intensity, even at roughly similar income levels. Therefore, this finding may be useful when interpreting the results if households with the same level income exhibit different energy consumption behavior.

Brandon and Lewis (1999) identified the number of occupants as an important determinant of energy use behavior, as energy use increased with increasing household size. Gatersleben et al. (2002) also found that lager households used more energy compared with smaller households.

2.2.2 Housing Characteristics (Designed environment)

House size and the age of the house could be other variables that may influence the household's energy conservation behavior. Johnson-Carroll et al. (1987) found that as house size increased energy consumption also increased. Yet, Wilhite & Ling (1995) found there is no significant association between house size and energy consumption. Usually, older homes are less energy efficiency because of poor insulation and design (Brandon & Lewis, 1999). Therefore, older homes may be associated with energy inefficiency problems that can be solved by upgrades and renovations of the homes. Examples might include upgrading the home's insulation, changing to a more efficient furnace, and purchasing energy efficient appliances, heating or air conditioning systems.

2.2.3 Occupant Attitudes (Social Environment)

Reviewing and understanding public attitudes is crucial to the successful development of energy efficiency programs, since promoting energy efficiency improvements requires involvement of one or more levels of end users, of which the public in its role as consumers is very important. Without public support and understanding, improving energy efficiency will be difficult if not impossible. As early as the 1970's, Canadians began to be concerned about energy as a natural resource. Energy policy in Canada has focused on sustainable energy development (Hageletam & Hollins, 2002). In the 1990s, the idea of bringing in citizens and other stakeholders in the policy-making process began to emerge (Hageletam & Hollins, 2002).

Guerin et al. (2000) indicate that the occupant attitudes and beliefs influence their energy consumption. Poortinga, Steg, & Vlek, (2004) point out that attitude is believed to be strongly influence by an individual's values. Values are usually conceptualized as important life goals or as normative standards that serve as a guiding principle in life (Poortinga et al., 2004; Rokeach, 1973).

Farhar & Houston (1999) were doing research about people's willingness to pay for efficiency and renewable electricity. In their study, they began to collect several national polls, and ask question like 'what have polls shown about the overall willingness to pay (WTP) for environmental protection and renewable electricity?' As a result, in several national polls, around 57% to 80% of the public said they would pay more for "electricity produced in a cleaner way" (Farhar & Houston, 1999; Farthar, 1994). However, when it comes to spending real money, the percentage of the public that are willing to pay extra for energy from environmentally friendly sources declined (Hageletam & Hollins, 2002; Farhar & Houston, 1999; Schulze, 1994). Based on this evidence, Farhar & Houston (1999) conclude that the percentage willing to pay extra for green energy tends to be higher in surveys than in a real situation. They suggest that income and education will made a difference in WTP as well.

A similar study by Zarnikau (2003) suggests that the willingness to pay a premium for renewable energy and energy efficiency among the general public can be explained by factors such as age, income, and education. He also points to increased public interest in energy efficiency as a primary energy option as well as increased interest in renewable energy as a first choice resource option. He states that as the public became better informed about different energy source options, they may adjust their interest in energy efficiency and renewable. The findings show that more available information about energy resource options will increase the public's willingness to pay extra for both renewable energy and energy efficiency. The findings also indicate that more respondents were willing to pay a small extra cost, whereas willingness to pay large extra cost decreased (Zarnikau, 2003).

There are two occupants' attitude variables that may influence energy conservation. The first is the desire for comfort. As early as 1978, Seligman et al. (1978) found that the desire for comfort is a strong predictor of household energy use behavior. Guerin et al. (2000) concluded that comfort was the most important attitude and suggested that attitudes involving comfort have more influence on conservation behavior than conservation attitudes. The other variable is health concerns which have been identified as being influential in determining household energy use behavior (Seligman et al., 1978).

Rowlands, Parker, & Scott (2001) conducted a case study in Ontario Canada about citizen and consumer attitudes towards electricity industry restructuring. The study focused on Waterloo Region which is a community in southwestern Ontario with a population of 450,000. They gathered public opinion about issues related to energy efficiency and use, energy policy and global climate change as these relate to energy restructuring concerns. They sent out 594 questionnaires to the region and got 386 questionnaires back. Therefore they have 65% response rate. By using these respondents' results to compare the Waterloo Region, Ontario or even Canada, the results shows the household income was a lot of higher than other areas and a higher proportion of respondents had a university degree or higher education. Thus the results of this survey may not be comparable with results in other areas. Rowlands et al. (2001) created several questions to examine citizens' attitudes towards green energy sources. The results show that 8% think the government should subsidize power generation from "Green" sources and 50% of respondents thought everyone should contribute through slightly higher electricity rates. Only 16% of respondents said green power should not be built (Rowlands et al., 2001). These results suggest that the majority of respondents are pro-environment and may be willing to support energy conservation and renewable energy initiatives.

Rowlands et al. (2001) also create some questions for consumer side. Respondents who had chosen a company to be their electrical energy provider were asked to rank factors such as price of the electricity, quality of customer service, reputation of the company or reliability of electricity in influencing their decision. Approximately 69% of respondents gave price and reliability high priority, putting these two factors ahead of environmental concerns. Another 31% of respondents selected reputation of the electricity provider as their first or second most important factor. From this analysis, they concluded that most consumers would demand low prices for electricity.

From Rowlands's et al. (2001) study, it can be seen that the motivation to conserve may influence the energy conservation behavior. Gmelch & Dillman (1988) identified and ranked four motivators of household energy conservation. The first one was economic benefit (reduce energy bills), the second one was conservation ethic that refer to the feeling of doing one's share by reducing usage, the third one was personal benefit means maintaining personal comfort and the value of homes, the last one was social conformity refer to peer pressure and guilt for not conserving. The first two motivators were a lot of stronger than personal benefit and social conformity (Gmelch & Dillman 1988).

Stern (1999) states that "…environmental degradation comes from economic activity, that consumer expenditures account for most of gross domestic product , and therefore changing consumer behavior can go a long way toward reducing environmental degradation." This statement helps researchers to understand that changing consumer behavior can make an important difference for environment, and conducting consumer research can help policy markers understand this behavior and thus develop policies that are more effective in influencing consumer behavior (Stern, 1999). Policies may seek to change consumer behavior by providing new and beneficial technology, changing financial and other material incentives, changing attitudes and beliefs with education and other information delivery tools (Stern, 1999). In the behavior domain, Stern said, a large number of individual behaviors may change the biophysical environment in direct or indirect ways, and it is important to remember that consumer behavior is highly situational.

2.3 Barriers to Participation to Energy Efficiency Programs

In order to successfully promote energy efficiency programs, it is important to consider what barriers or obstacles may inhibit improvement in energy efficiency. If the barriers can be understood, programs and policies can be modified to encourage wider participation now or in the near future. By understanding those barriers, the factors that may influence participation also can be explored.

As early as 1982, Anderson wrote a paper about barriers to consumer choice of energy efficient products. He identified three major barriers to energy efficiency choice. The first one is limited cognitive capacity, which means consumer' limited cognitive capacity will increase the difficulty of attracting attention to energy efficiency information and usage of energy information (Anderson, 1982). The second barrier is salience of energy information. Anderson (1982) gives an example based on manual defrost versus frost-free refrigerator models: there are only small differences in operating costs for refrigerators of approximately the same size. For example, if a consumer purchases a 15 or 16 cubic foot refrigerator, then he can save 10 Canadian dollars a year. But he could save 26 Canadian dollars a year by buying a 13 cubic foot refrigerator. As a result, consumers are very slow in adopting energy consumption as a salient attribute (Anderson, 1982). Dominance of retail sales staff was noted as the third barrier. Because of their dominant role, salespeople in a store can help with consumers appliance purchases and can deliver information to the consumer about energy efficiency. Sales staff support is necessary to promote efficient products (Anderson, 1982; Day, 1976).

There is another paper by Reddy, which is more detailed and powerful than Anderson's research. Reddy (1991), explored the barriers to energy efficiency at the consumer level and suggested that several types of consumers act as barriers including the ignorant, the poor / first-cost-sensitive, the indifferent, the helpless, the uncertain and the inheritors of inefficiency. In this paper, the author tried to find out the origin of these barriers and to suggest methods that either by themselves or combination with other methods could overcome these barriers. Consumer support for and participation in energy efficiency programs depends on the consumer knowing about energy technology (Reddy, 1991). The easy way to solve this barrier of ignorance is to provide broader information in different ways, such as door-to-door canvassing, leaflets through the mail, newspaper, magazines, radio and television (Reddy, 1991). Both quality of information and effectiveness of the communication are important. He also mentioned information, education and training are mainly important for energy intensive industries. The barrier of ignorance can be overcome by the effective provision of relevant high quality information, and education and training of the consumer. This is also one factor that may influence consumers' decisions to participate in action for energy efficiency.

The second barrier in his study is the first-cost-sensitivity which is particularly an issue for low income households. Consumers with full knowledge of the benefits of energy efficiency improvement may avoid investment in efficient energy devices and equipment because of their higher initial cost (Reddy, 1991). Therefore, the consumer will ask questions like: "do the energy savings and other benefits justify increased investment on the efficiency improvement?" (Reddy, 1991) He said the answer for this question really depends on whether the consumer would like to invest capital resources now and get the regular benefit of lower energy costs in the future. That also means these consumers would like to minimize the life-cycle costs instead of minimize first cost. In his paper he uses the term 'pay back period' which is the time required to recoup an energy investment through the money saving from the energy bill and the consumer discount rate (CDR). From information presented in the paper, as the income of the consumer increases the CDR used for investment decisions will decrease. In the end, Reddy (1991) concludes that "innovation financing involving the conversion of first costs into operating costs is a crucial method" to help overcome the second barrier.

The third barrier is indifference. This involved consumers who had no problem with the first two barriers but who were indifferent to energy efficiency. This may be due to energy savings for these consumers not being significant enough in relation to their total expenditures (Reddy, 1991). Reddy (1991) suggests government intervention through regulations, standards, labels, restriction in supply, etc., alone or in combination can overcome the indifference barrier.

Helplessness is the fourth barrier. This class of consumers needs help in identifying, procuring, installing, operating and maintaining energy efficiency devices and equipment and for whom the required assistance may not have been provided (Reddy, 1991). Thus, an efficiency improvement industry must be considered to provide help. The last barrier is inheritors of inefficiency. The most common situation for this group is people who rent a house that is energy-inefficient or has appliances that are inefficient (Reddy, 1991). This may be partly solved by labeling equipment with energy performance data, which may put pressure equipment providers to make efforts on energy efficiency products (Reddy, 1991).

Meyers (1998) described six different classes of barriers to improved energy efficiency which are macro-economic conditions; energy pricing; international flows of capital, technology, and knowledge; institutional weaknesses; market behavior and features; and features of energy efficiency products or services. His study focused on developing and transitioning countries. He mentioned the above barriers occur in all countries. However, the first four classes of barriers play an important role in developing countries. The first class of barrier is caused by different factors such as low level of competition among firms due to regulation of domestic market and/or policies that forbid foreign products entry to the market and as a result this barrier impacts both demand and supply sides of the market; high level of income inequality means poor households are least able to purchase higher cost energy-efficient technologies, and they discount future savings more than wealthier households (Meyers, 1998).

The second class of barrier is related to energy prices. Meyers (1998) said “a problem related to energy pricing is the existence of weak feedback between energy consumption and payment for energy”. The third barrier is related to international flows of capital, technology, and knowledge. The large fluctuations in exchange rates will bring problems for capital investment, and usually small market size limits the ability of countries to gain access to new technologies and lack of reliable resource and information will be obstacles to knowledge gathering (Meyers, 1998).

The fourth barrier is related to the weakness of education, research, government and financial institutions. Advanced education institutions are inadequate in most developing countries and lack facilities for research. Government institutions lack trained personnel to design and implement energy efficiency programs, and have difficulty attracting or retaining the staff to work on this area because the salaries are lower than in private sector. Financial institutions lack experience with evaluating investments associated with energy efficiency (Meyers, 1998).

The fifth barrier addressed in Meyers (1998) study is market behavior and features both in demand side and supply side of the market. For the demand side, the barrier is the lack of information: consumers and mangers often lack information about the costs and benefits of higher energy efficiency technologies or services. While the barrier for supply side of the market can refer to weak marketing capabilities of suppliers, for example, companies may lack the skills for sufficient marketing of more efficient products or services.

The last barrier in Meyers (1998) study related to features of energy efficiency products or services. New or unfamiliar technologies introduce performance uncertainties which can be minimized by acquiring better information (Meyers, 1998). High first cost is another common feature for energy efficiency technologies. Furthermore, the inseparability of energy efficiency from other product features may increase the total cost of a product beyond the ability of the consumer to pay.

Brown (2001) addressed three market barriers for energy efficiency improvement. First, he pointed out that energy efficiency is not paid attention to by most consumers, because energy costs are lower than costs of other goods and services. Therefore, there is a barrier due to the low priority of energy issues and opportunities for energy efficiency. This in turn reduces producer interest in providing energy efficient products (Brown, 2001). The second barrier is capital market which can inhibit energy efficient purchases (Brown, 2001). The last serious obstacle in Brown's (2001) study is incomplete markets for energy efficiency.

More recently, Stewart & Fry (2006) developed a paper based on low income households. They argued that low income consumers face significant barriers to participation in energy efficiency programs, and these barriers are also factors that influence participation rates. These barriers include low income households' lack of ability to invest in better insulation or new appliances, which prevents them from reducing energy consumption significantly. Many low income households are tenants who do not own the building or appliances and can't benefit from long term investment in more energy efficient appliances or building retrofits. Furthermore, there is no financial incentive for landlords to invest in energy efficient equipment because the landlord can pass higher energy costs to the tenant (Stewart & Fry, 2006).

2.4 Summary

Residences are being built in more energy efficient ways than a century ago, but many opportunities for improvement remain. By improving our understanding of government initiatives, public behavior and barriers to participation in energy efficiency, it may be possible to design more effective policy interventions and to explain their rationale to the public and get more people involved in energy efficiency programs.

Based on the first part of this chapter, it can be seen that federal and provincial government and local community all care about energy conservation and their environment. Federal government offers grants to local community to encourage participation in energy efficiency programs. Local communities also create their own energy efficiency programs through private organizations or have partnerships with government to deliver energy service and products to residences.

The literature review revealed several variables that may influence energy conservation, such as age , education, household size, household income, or physical house size were more or less related to energy consumption behavior. Some barriers have been identified in order to improve the outreach to the public and also help to create different strategies for success in energy efficiency program.

Chapter 3 Spatial Analysis using GIS

3.1 Introduction

This chapter will focus on a comprehensive GIS analysis of the Waterloo Region to identify physical and socio-demographic characteristics of neighborhoods that may affect participation rates in the REEP. GIS analysis was the first method adopted in this research. The methodology discussed in this chapter was used to analyze and present the results about the spatial distribution of participating households in the REEP from 1999 to 2006. Finally, a series of thematic maps were used to assess the importance of factors which may affect participation rates.

3.1.1 Description of the Study Area

The Regional Municipality of Waterloo was selected as the study area for this research. The Waterloo Region was created in 1973 from the County of Waterloo and a small section of the County of Wentworth. Originally, there were fifteen local governments which were reorganized into seven Area Municipalities (Regional Municipality of Waterloo, 1998). Waterloo Region is a community of approximately 506,800 people and 149,269 eligible households at the end of year 2006. The region is situated in southwestern Ontario (Canada), located 100 kilometers west of Toronto and comprised of three cities (Kitchener, Waterloo, and Cambridge), and four rural townships (North Dumfries, Wellesley, Wilmot and Woolwich) (see Figure 3.1). The percentage population change between 2001 and 2006 in Waterloo Region was 9% (Statistic Canada, 2007). Waterloo Region covers a land area of 1,369 square kilometers. This is one of Canada's fastest growing communities and is expected to reach 729,000 people by 2031. The region has twice the growth rate compared to the national average. The median population age is one of the lowest in Canada (Region of Waterloo, 2008). The cities within Waterloo Region maintain an extensive network of street, community trails, parks, and green space. The region is well served by public transport, by train, plane and bus.

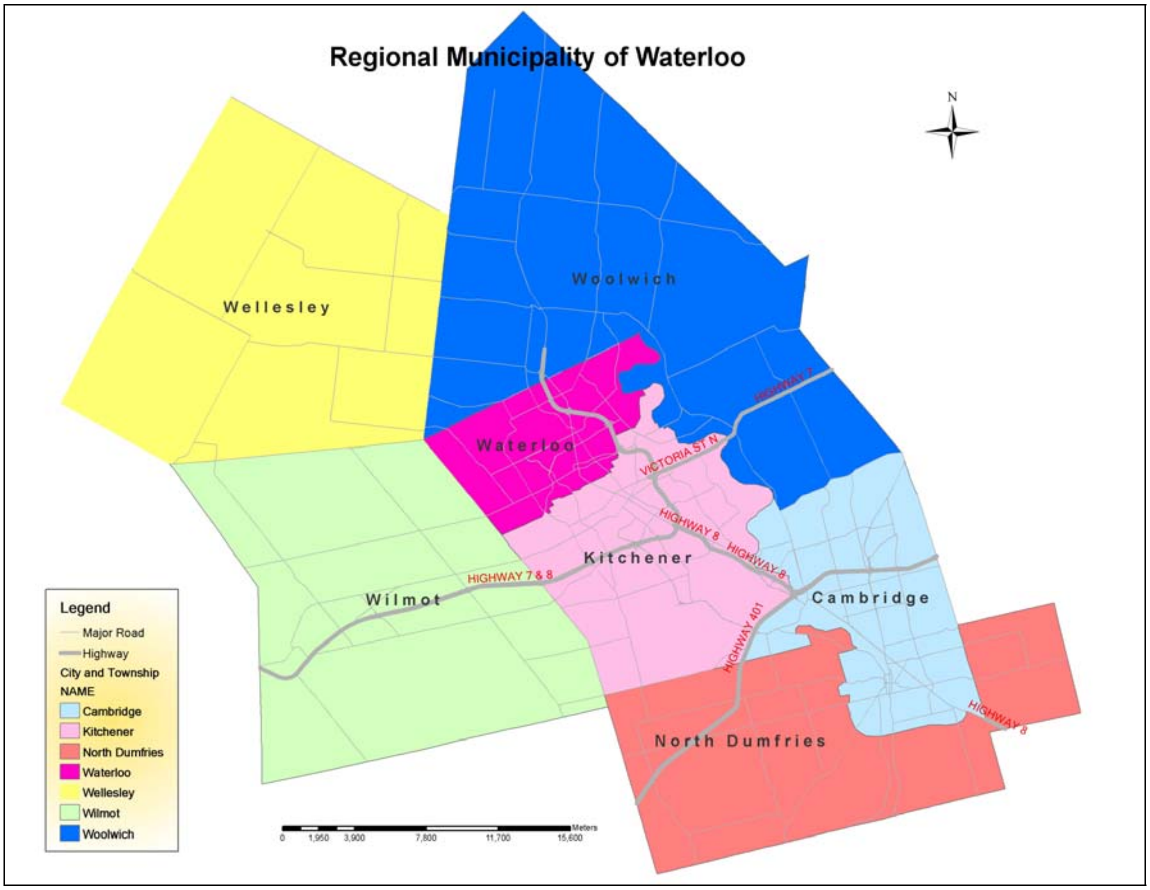


Figure 3.1 Waterloo Region

3.1.2 Analysis Goals and Objectives

The primary goal of this study is to assess the importance of factors that may influence participation in the REEP. Understanding the factors that influence households' decisions to participate in the REEP may make it possible to develop marketing campaigns that will encourage greater participation in this kind of program in the future. To determine which factors may affect the REEP participation rate, two spatial statistics methods were used. These will be discussed in chapters 4 and 5. In this chapter, a Geographic Information System (GIS) was utilized from both a spatial analysis and a visualization perspective to analyze the spatial distribution of participating households in the REEP and visually assess physical and socio-demographic characteristics that may influence participation rates.

There are three research objectives about the REEP that will be explored in this chapter. First, to examine the evolution of the spatial distribution of participating households from 1999 to 2006; second, to explore which sources of information (marketing method) played an important role; and finally, to decide which factors may affect the REEP participation rate.

3.2 Data Preparation

Before considering the methods used in doing the analysis, the following sections describe the data sources used in this research and the pre-processing required to build the GIS database that was used for visualization and spatial analysis.

The spatial and attribute data used in this thesis are secondary data, retrieved from Statistics Canada, the REEP office and the University of Waterloo Map Library. The specific data sets used to develop the analysis in this chapter are listed in Table 3.1. The REEP data consists of two datasets, called REEP Data and REEP Follow-up Data. The first dataset includes all participating households, and the second includes participating households who also completed a second follow-up evaluation after completing upgrades.

3.2.1 REEP Data

REEP was founded in May, 1999 and faced the sudden termination of its primary service (residential energy evaluations) in April 2006 when the federal government cancelled the EGH program and associated delivery contracts. A revised program, ecoENERGY, was launched in 2007, but 2006 makes a natural break in the data for study purposes.

The main service provided by REEP was an evaluation of the energy efficiency of houses. The federal EGH program was based on the Hot2XP computer model of house attributes and energy performance. This program can help home and property owners learn how to save energy, reduce energy bills, as well as improve home comfort. To take full advantage of this program, households were required to complete an initial evaluation of their house and a second follow up evaluation after they had completed recommended upgrades. Once the home owner books an appointment for home energy efficiency evaluation, a Certified Energy Advisor will come to that home to determine the home's energy efficiency performance, provide a rating on a scale of 0-100 and identify sources of heat and energy loss (REEP, 2007a). After the evaluation is complete the Advisor will provide a detailed report to the homeowner containing the recommended parts in the home that need upgrading, energy savings estimate after upgrades, and an *EnerGuide for Houses* label and rating from 0 to100 that represents the level of efficiency of the home (see Appendix E). Also, if participants would like to be eligible for government grants, they have to do a second evaluation (follow up evaluation) within 18 months and submit it to NRCan. The amount of grant money available to households is based on what kind of improvements has been made to the home, and the criteria can be found in Natural Resources Canada (2008). Sometimes participants may take longer than this time to do the retrofits of their homes, and they will still consider as eligible dwellings.

3.2.1.1 REEP Data

The REEP data includes all participating households and their energy consumption information, contact information, estimates of expected energy savings from upgrades, and information on how they became aware of this program (source of information).

3.2.1.2 REEP Follow-up data

The REEP Follow-Up data has the same attributes as the REEP data and also includes information about actual energy savings for each household that participated in the second evaluation. The Follow-Up data includes a subset of households from the REEP data, with each household being identified in both datasets by a unique NRCanID. This allowed the analyst to link all the information together. The REEP data were used to calculate the potential energy savings by joining this program, and the REEP Follow-Up Data were used to calculate the actual energy savings achieved by implementing the advice provided by this program.

3.2.2 Other Data

3.2.2.1 Census data

Census data was used to integrate into the GIS data as an asset to analyze the relationships between participating households and explanatory variables. The CDAs include information about census counts for Waterloo Region such as dwellings (households), population with higher education, population over 65, average household income, owner occupied, and dwellings built before 1970. For some parts of the analysis, census variables at the CDA level were converted into percentages.

The use of census data was necessary because the REEP database did not include demographic information about participating households. This introduces the problem of ecological fallacy, in which inferences about individual behavior are made based on aggregate data. This problem is discussed in more detail in chapter 5.

3.2.2.2 GIS Data

The Waterloo Region Address Point shapefile was used in conjunction with Teranet property parcels data to locate participants within the Waterloo Region. A geocoding process was required before spatial analysis could be performed. Street network layer and boundary layer helped identify the accurate location for the participating households.

All data discussed in Table 3.1can be integrated together using some spatial and non-spatial database operations, such as select, join and spatial join (after geocode), at the CDAs (census dissemination areas) level. When this research started, the 2006 Census of Canada data were not fully released, therefore, all the data from the CDA level used in this thesis are from the 2001 Census of Canada for the Waterloo Region.

Table 3.1 Data Summary

Data Category	Contents	Data Source	Object Type	Time	Description
REEP Data	Contact	REEP Office	Table	1999-2006	Basic information about participants (address, postcode)
	Energy consumption and Energy equipment	REEP Office	Table	1999-2006	Information such as annual fuel consumption, space energy consumption and furnace type, furnace efficiency, A rating and so on
	Estimate Energy saving by REEP	REEP Office	Table	1999-2006	Energy saving information about all energy consumption and U rating
	Household Characters	REEP Office	Table	1999-2006	Information such as year of construction, house size, footprint, ownership, house type.
	Evaluation source	REEP Office	Table	1999-2006	How did clients hear about REEP (coded by number)
REEP Follow-Up Data	Contact	REEP Office	Table	1999-2006	Basic information about participants (address, postcode)
	Household Characters	REEP Office	Table	1999-2006	Information such as year of, house size, footprint, ownerships, house type.
	Energy consumption and Energy equipment	REEP Office	Table	1999-2006	Information such as fuel consumption, space energy consumption and furnace type, furnace efficiency, A rating and so on
	Exactly Energy saving by REEP	REEP Office	Table	1999-2006	Energy saving information about all energy consumption and U rating
GIS Data	Waterloo Region Address Point	UW Map Library	Point	Retrieved from year 2007	Contain all the information about the addresses in Waterloo Region
	Street Network	UW Map Library	Polyline	2001	Street name
	City Boundary	UW Map Library	Polygon	2007	Name of cities and townships
Census data	FSA (forward sortation area)	Stat. Canada	Table	2001	Information includes 3 digital postcode
	CDA (census dissemination area)	Stat. Canada	Table	2001	Information includes education, income, gender, construction period, and other household characters.

3.2.3 Pre-processing

3.2.3.1 Geocoding

Before spatial analysis could be performed, a geocoding process was used to create point features representing REEP participant household locations. Geocoding is a process for importing non-map data into a GIS using a locator service containing address information. For this study, a locator service was created by spatial joining address points obtained from the Regional Municipality of Waterloo to property parcel polygons obtained from Teranet. This creates a new version of the parcel polygons feature class in which each polygon has an address attribute, allowing it to be used as the reference layer for a locator service. In the REEP database, participating households were identified by their street address. In the geocoding process, the parcel polygons reference layer is searched for parcel polygons that have addresses that match the addresses of REEP participants and the centroid of the matching polygon is returned as the location of the participating household. Thus, this process allowed the REEP participant households to be located on the map of Waterloo Region. Figure 3.2 below shows the general steps taken to geocode the REEP participant addresses for all the participants in the REEP Dataset and the REEP Follow Up Dataset.

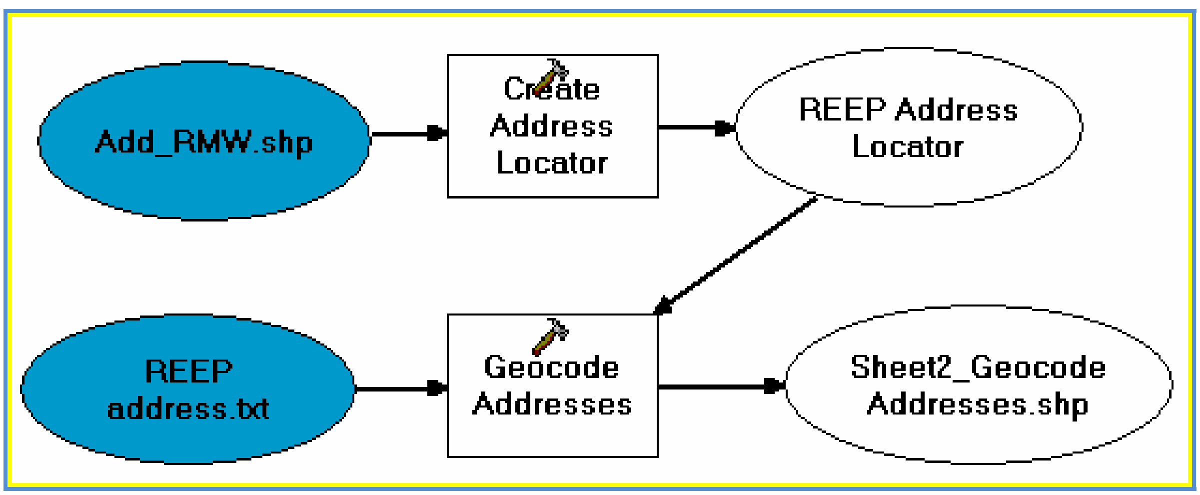


Figure 3.2 The General Steps Taken to Geocode the REEP Participation Address

After the geocoding was completed, the gecoding result (new address point layer) was added to the map. The accuracy of the geocoding process can be assessed by determining the number of addresses that were correctly matched. Under the 'tools' menu, the ***review/rematch addresses*** tool was used to interactively match unmatched addresses. The FSA data was used to eliminate households that were not in the Waterloo Region. Then by using an orthophoto image of Waterloo Region, the GIS Locator provided by the Waterloo Region GIS department, it was possible to locate most of the unmatched addresses.

Table 3.2 shows the statistics from the automatic and interactive address matching. From this table, it can be seen that the automatic matching process correctly matched 85% of the households participating in REEP (6132 addresses) leaving 15% (1082 addresses) unmatched. For the REEP Follow-Up data, 83% (1051 addresses out of 1266 addresses) matched leaving 17% (215 addresses) unmatched. The quality of matching score can be improved by match interactively. The interactive address matching process presents the user with a list of candidate addresses for each unmatched address. All of the candidates of the unmatched addresses were evaluated and the best one was chosen. The candidates of the geocoded points were compared to where they were on the 'orthophoto image layer'and the 'Waterloo Region GIS address locator' and the closest candidate was chosen as the matching point. The map from GIS locator was used to find the proper street name, house number, or township name in order to accurately find the location of the address point. There were quite a few address points that had to be matched in this way. Therefore, the geocoding was not perfect. However, it was not an overwhelming number, so the geocoding can still be considered satisfactory. For the REEP data, 99% (7118 addresses out of 7214 addresses) were matched leaving 1% (96 addresses) unmatched, and all 99% out of 100% matched had a matching score greater than 80. Some of the addresses however did not get matched because they didn't exist or it was too difficult to accurately assess where the location was or had another limitation that will be discussed later in the data limitation section. In conclusion, the overall matched percentage has been increased 14% and 16% for REEP data and REEP Follow-Up data respectively.

Table 3.2 Geocoding Result for the REEP Data and the REEP Follow-Up Data

	Match Automatically		Match Interactively	
	Matched (%)	Unmatched (%)	Matched (%)	Unmatched (%)
REEP Data	85% (6132)	15% (1082)	99% (7118)	1% (96)
REEP Follow-Up data	83% (1051)	17% (215)	99% (1249)	1% (17)

* The number in brackets is the number of addresses

3.2.3.2 Data Limitations

Several problems were encountered when geocoding the REEP household address data. The following subsection will discuss these problems.

Geocoding

Many of the problems involved the difficulty in evaluating the accuracy of the available candidates based on the resources used as a guide for matching the points interactively. Candidate addresses were compared to other material in order to select the best match. However, sometimes none of the candidates appeared to be suitable and no choice was better than any other. For example, in the REEP data sets, there is an address 274 Lancaster Street West, Kitchener which should be a residence, but there is no candidate matching this house number (Figure 3.3). The Region of Waterloo GIS Locator was used to check for a suitable candidate. Again there was no matching address and the two closest street numbers were commercial rather than residential properties (Figure 3.4). In some instances, the interactive matching process failed to suggest any candidates.

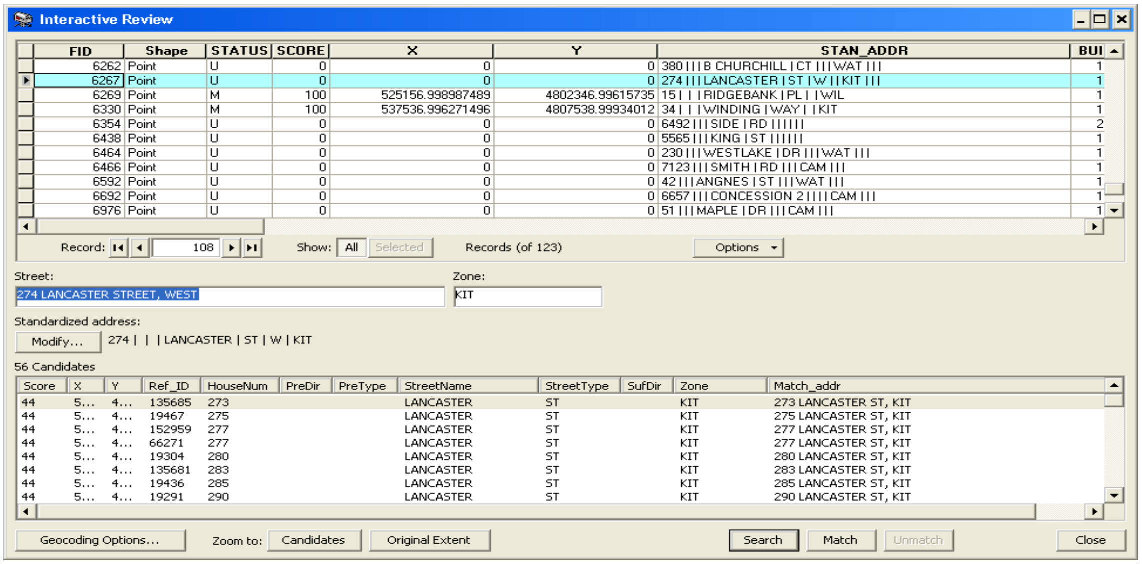


Figure 3.3 Candidates for 274 Lancaster Street West, Kitchener

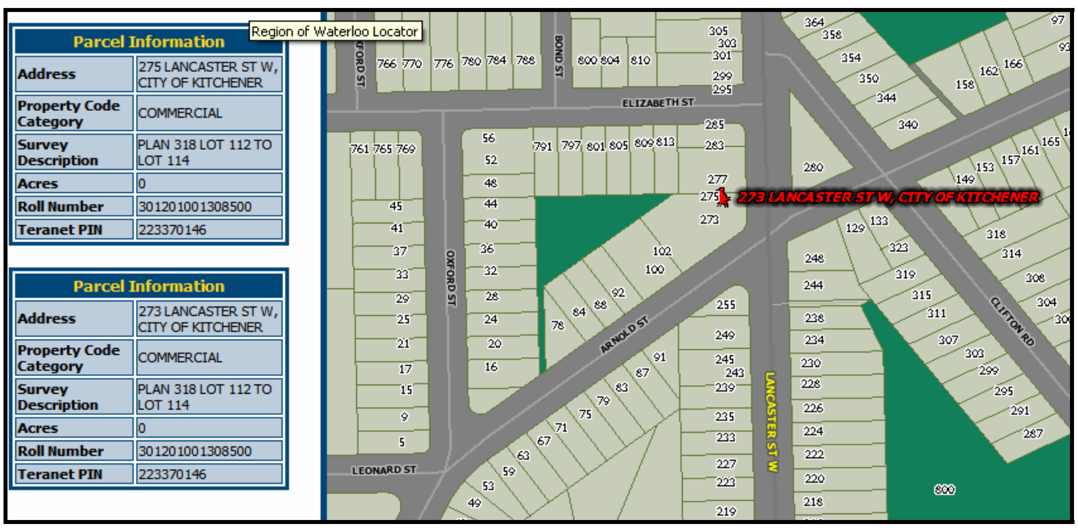


Figure 3.4 Parcel Information

The second problem is spelling errors in either the REEP addresses or the locator service that may result in low match scores. Spelling sensitivity controls how well the addresses from the REEP_address.xls must match with the addresses of Add_RMW.shp. For example, if 278 Hazle St. (correct spelling is Hazel St.) was to be geocoded with a spelling sensitivity of 70 or 90, it won’t get matched. However, if the spelling sensitivity was 20, 278 Hazle St. may have chance to be matched, although the spelling is not correct.

The third problem is that some names of city zones in the REEP data were different from names used in the address point shapefile. For example, ELMIRA is in Woolwich township, so if the zone name is changed to WOO (short term), this address can be matched.

The fourth limitation is that some street names were not available on the address points layer. These may be new streets or streets that have changed to a new street name. This situation made it more taxing to properly match the points. Extra care had to be used to ensure that the right points on the right streets were being looked at and matched.

The last thing is the sources themselves that were used to help match the points may also contain flaws. Although this may not be a common occurrence, it may be possible. Therefore, the matched geocoded points may contain the same inaccuracy.

REEP Data

There may also be problems with the REEP data sets that might affect the geocoding results. During the data entry process, the staff may enter wrong street name, house number, city's name, or participant's postal code. Those data entry errors are going to cause problems during the geocoding procedure. For example, there are some house numbers that can't be found in the address locator, either the house number is a lot of bigger than the available candidates or smaller than candidates. Some records don't have a house number in front of addresses, or just have a mailing route for mail delivery, not a detailed address for the household.

There was another problem in original REEP data sets. In the FSA, there is a postcode N3A, however, there is no information for N3A for FSA. Based on information was provided by UW Map library the N3A that it didn't exist in 2001. The N3A indicates an area in Wilmot, and all of Wilmot was in N0B in 2001 but became N3A in 2003. In order to make sure the data postcode can be changed to N0B from N3A, the year of evaluation must be checked. By checking the original file for year of evaluation it was found that all N3As are from year 2003 to 2006, so that means the N3A participations can be added into N0B, to calculate the participation rate.

Also there were around 30 postcodes with typing errors, these errors were corrected by using Google Map, basically just type the participant address and the corresponding postcode was returned and could be checked to see whether there was a mistake or not. Since the study area was limited to the Region of Waterloo, addresses outside the Waterloo Region were excluded. At the end, there was only one participant address that could not be found in this area.

3.2.3.3 Participation rate calculation

The final geocoded REEP dataset included 7118 households that could be reliably located within the Waterloo Region. A spatial join was used to join all participant addresses to the CDAs, thus assigning a unique CDA identifier (DAUID) to each REEP household. A summary table was generated containing counts of the number of REEP households in each CDA. Apartments were not eligible for REEP, thus when the census data were used to calculate the participation rate, the number of apartments was excluded from the total number of dwellings. The REEP Participation Rate was calculated as:

REEP Participation Rate = number of household participants in REEP/ (total number of households – number of households in apartments)

3.3 Methodology

In this section, three hypotheses have been developed that may help explain participation in REEP:

1. The spatial pattern of participation changes over time and tends to cluster in urban areas.

If the distribution of REEP participants changed over time and covered the entire region, that may signify the program was well developed; and if the cluster occurred in urban areas in global scale that may also indicate the program was well developed, since the urban areas have more eligible dwellings than rural areas.

2. The spatial distribution differs by source of information.

If the distribution differs by source of information that can help determine which marketing method played an important role.

3. Relationships exist between participation rate and explanatory variables.

If there is a relationship for any explanatory variable that can help to explain which groups of people have more potential to participate in this program. Based on the literature review and the data available from the census dissemination areas (CDAs), seven explanatory variables were hypothesized to influence participation rates (Table 3.3). Each explanatory variable represented a condition for what kinds of people were willing to join this program. In addition to the variables listed in Table 3.3, the percentage of population aged 6 years and under and urban / rural area were investigated in the analysis, however, these variables were found to be non-significant in relation to participating households in the REEP.

Table 3.3 Name and Description of Explanatory Variables

1. % of Population with Higher Education [a]
2. Average Household Income
3. Employment Rate
4. % of Dwellings Occupied by Owner
5. % of Population Aged 65 Years and Over
6. % of Dwellings Built Before 1970 (before energy crisis)
7. % of eligible households (dwellings) [b]

[a] attaining at least a university degree
[b] total number of dwellings excluding apartments

GIS plays an important role in addressing these issues. GIS technology allows integrating a wide range of social and demographic information in support of the analysis objectives. Firstly, the maps that need to be generated for this research were made possible because of the ability of GIS to display and operate data sets. The most important data sets were included as map layers: participating households, municipal boundaries, major roads and highways, and then the physical and socio-demographic information of CDA in Waterloo Region. The map layers allowed for visual observation of the spatial distribution of the REEP households from 1999 to 2006, as well as other information of interest.

In addition, correlation coefficients as a supplementary test were calculated in the statistical software program SPSS (SPSS, 2008) to assess the relationships between explanatory variables and participation in the REEP. The count value for each explanatory variable and number of eligible households (apartments were not eligible to participate) in each CDA were used as input in a bivariate correlation analysis to determine the strength and direction of the relationships

3.3.1 Map Design and Spatial Analysis

This section discusses the map design and spatial analyses in GIS and how to show the valuable results on each thematic map to explain the spatial patterns and spatial relationships between features. Three sets of thematic maps were designed in ArcMap using both REEP data and 2001CDA data to prepare the spatial analysis. The Steps involved in this analysis are provided below:

- Geocode participating REEP households
- Spatial join the cda2001 to the geocoded households
- Join the output spatial join table to cda2001 based on the common field DAUID, output layer call DA_ROW_UTM
- Add two new fields to the DA_ROW_UTM (create from cda 2001):

a) Number of participating households in each CDA

b) Eligible dwellings

c) Calculate participation rate

d) Calculate the % of population with higher education, % of dwellings occupied by owner, % of dwellings build before 1970, % of population aged 65 years and over, % of eligible households;

- First: map participating households based on geocoded households;
- Second: map participation rates based on DA_ROW_UTM;
- Third: map overall participation rates with hypothesized explanatory variables.

In addition, the detailed steps were provided in the data work flow diagram (see Figure 3.5 and 3.6). Also other relevant data in the study area helped to understand the spatial distribution, including city boundary layer, and street network layer. To illustrate the results on the maps, different symbology was used, including graduated colors, dots.

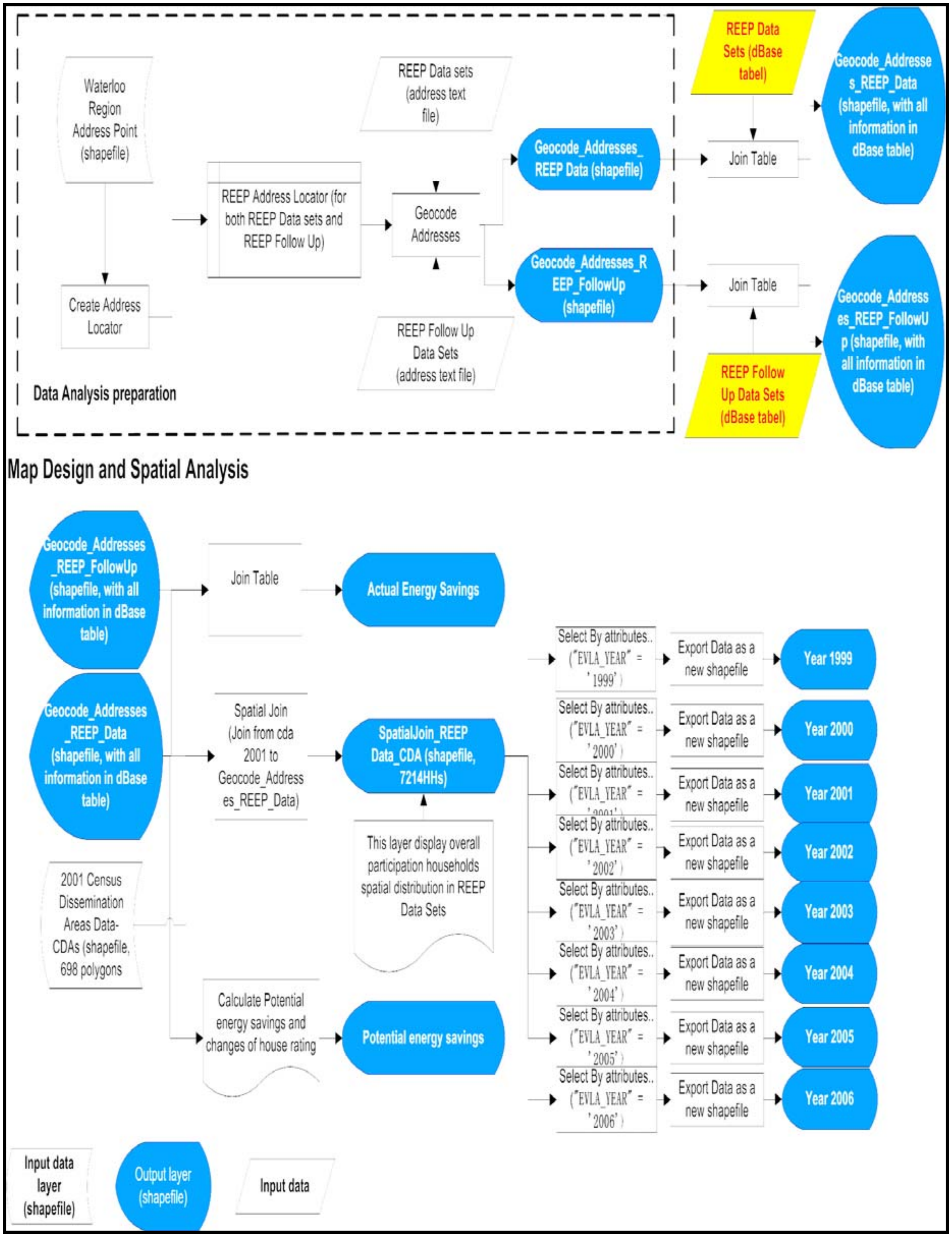


Figure 3.5 Map Design

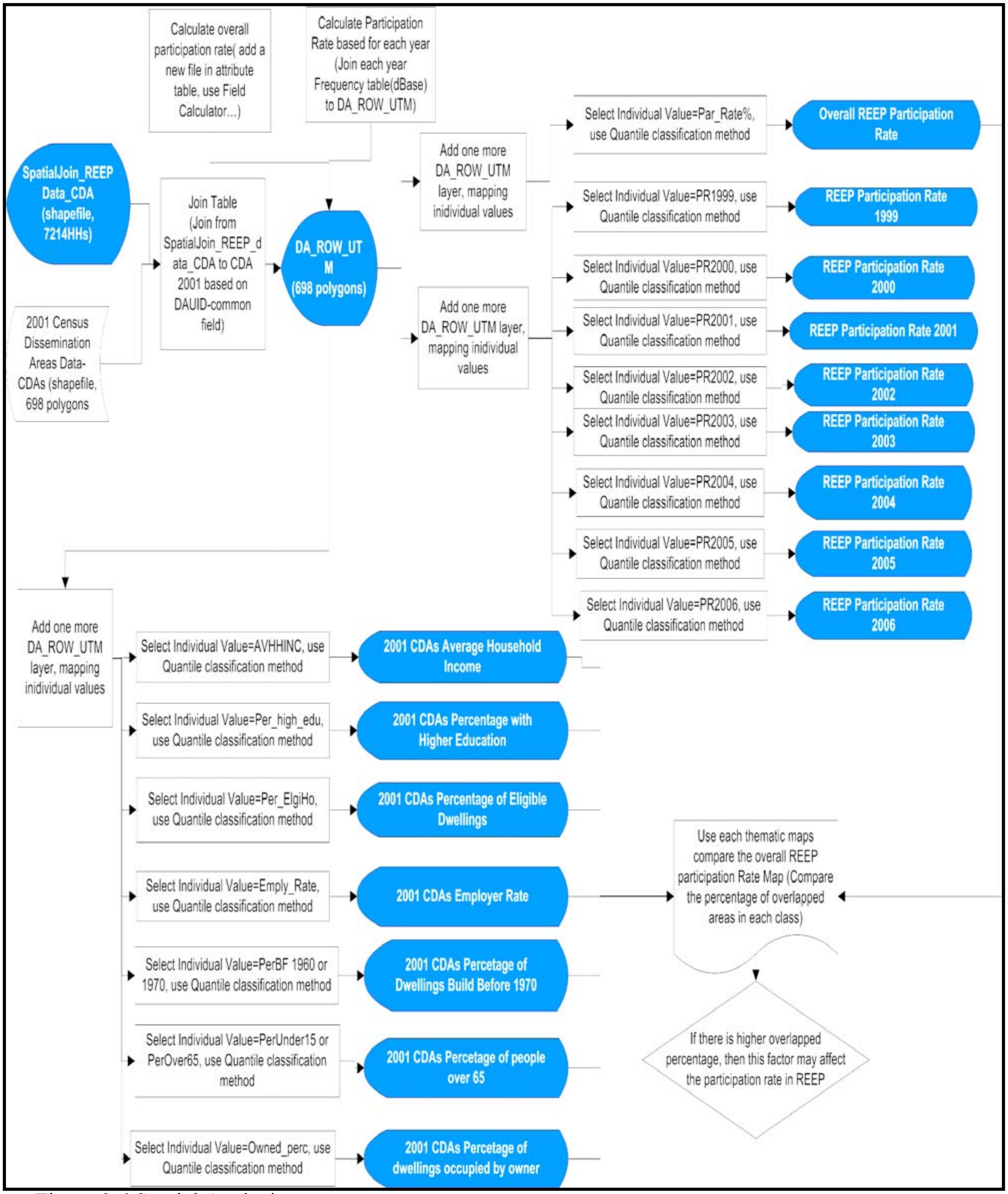


Figure 3.6 Spatial Analysis

3.4 Results and Discussions

This section presents the results of visual comparison of the spatial distribution of participating households and the spatial analysis. In the first part, the spatial distribution of participating households is mapped to allow comparison of the changes in spatial distribution from 1999 to 2006. In the second part, maps were prepared to show the evolution of participation rates. These maps include pie charts showing source of information in order to assess which source of information played an important role in different years. The third part of this section presents maps that compare participation rates with the distribution of each hypothesized explanatory variable to determine the relationships.

3.4.1 REEP participation Households Analysis

During the eight years (1999 to 2006), 7,214 households completed the evaluations in Waterloo Region (Figure 3.7). The overall spatial distribution of participating REEP Households is shown in Figure 3.8. From this figure, obviously, the participating households tend to cluster in urban areas, and other small communities like Elmira, Wellesley, Baden, New Hamburg, New Dundee, and Ayr; and tend to be scattered in rural areas. Figure 3.9 and 3.10 provide detailed information on how this pattern formed and how this program was developed in each year.

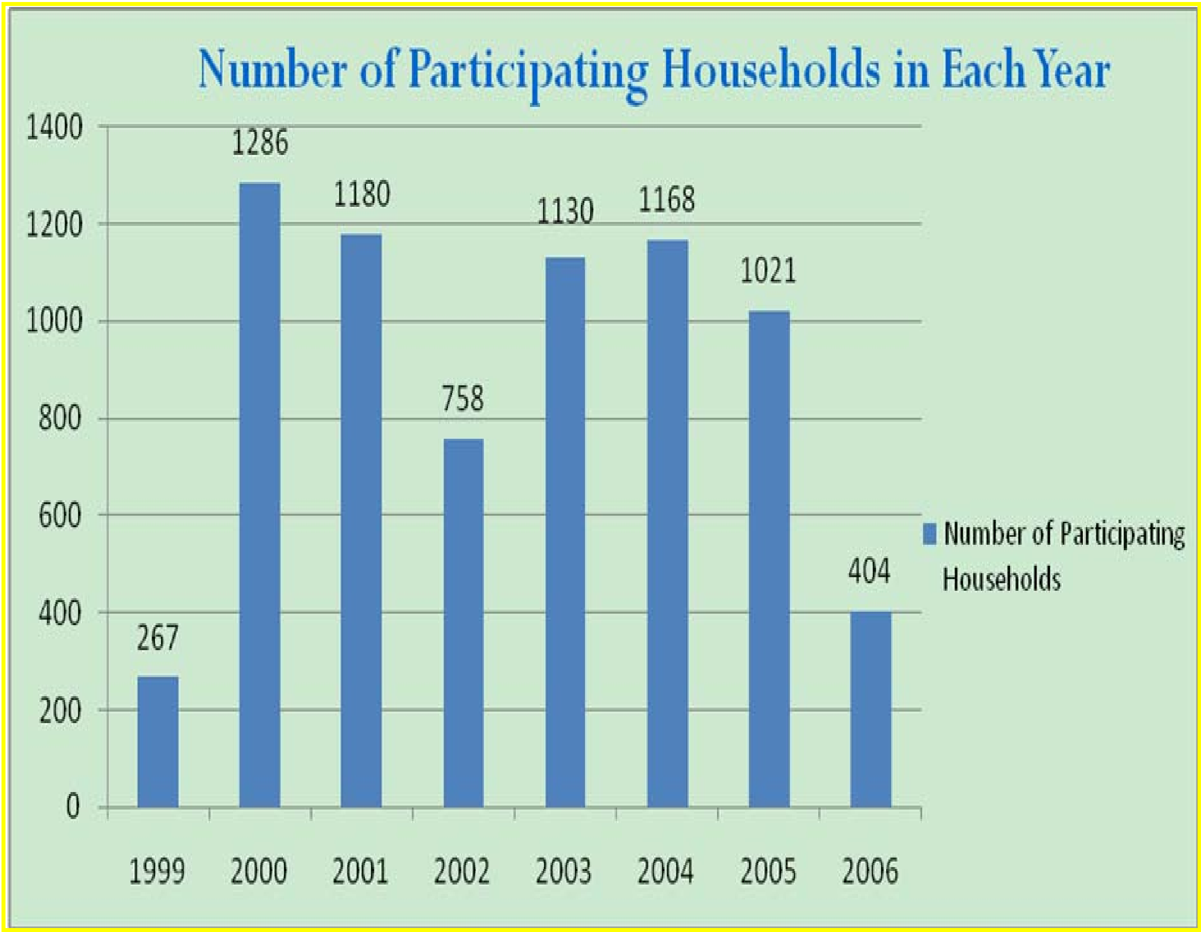


Figure 3.7 Number of Evaluations in Each Year

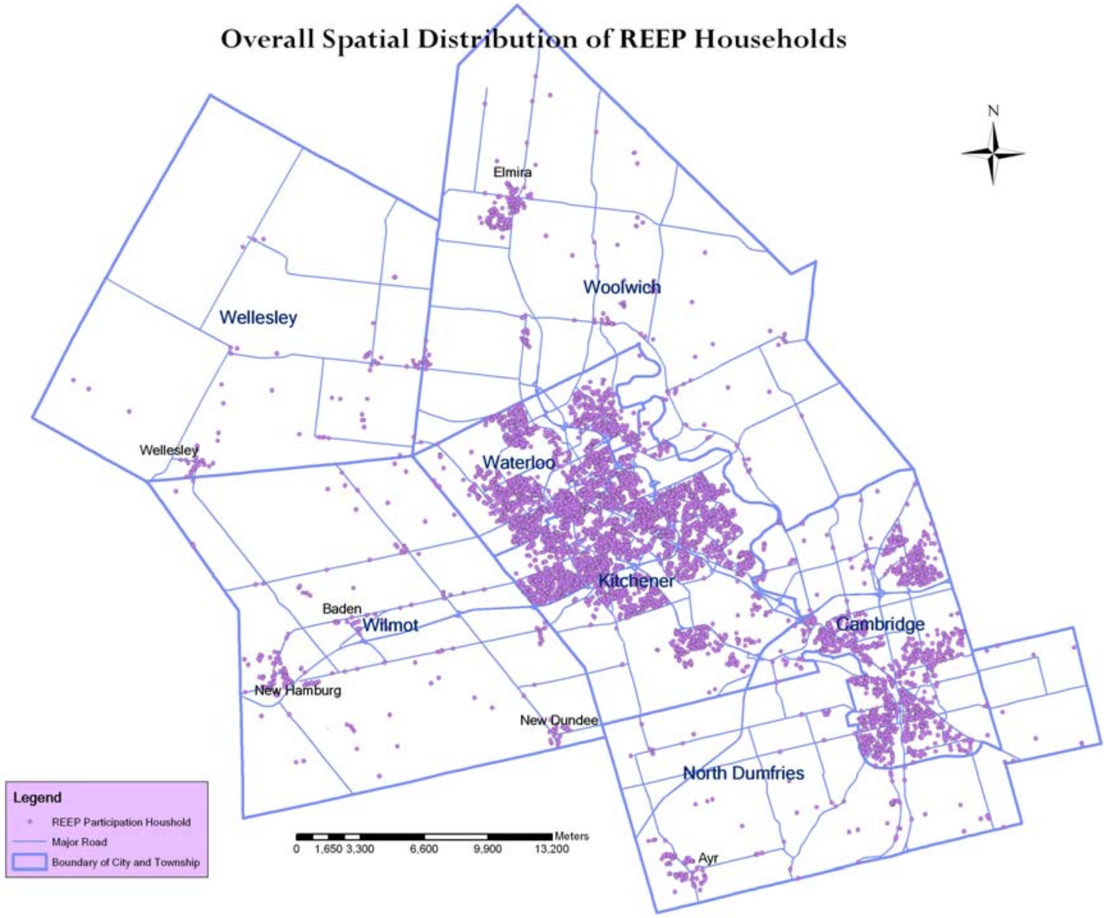


Figure 3.8 Overall Spatial Distribution of REEP Households

The REEP was started in May 1999, therefore, from Figure 3.6, it can be seen, that only 267 participating households in the whole region. Since this program was mainly founded by University of Waterloo faculty, initial participating households were close to the university and most were located in or near the City of Waterloo, with little participation by households in the smaller communities. By looking at Figure 3.9 and 3.10 from 2000 to 2006, the cluster tends to focus on cities more obviously. As the program developed, the participating households in rural areas tend to increase although they were still scattered. There is one explanation about this evidence. The cluster shows in urban areas because urban CDAs are much smaller than rural CDAs. There was a trend for the changes in the spatial distribution of participation. Participating households first developed from the City of Waterloo and City of Kitchener, and then expanded to the City of Cambridge, to form a very dense pattern through eight years. Furthermore participation seems to expand to southeast first, and then southwest, followed by the northwest and northeast areas.

Figure 3.7 shows the number of participating households in each year. There were few participants in the first and last years, in part because the program only ran for part of these years. However, 2002 also had a relatively small number of participating households. However, several reasons can be identified for the drop in the 2002 based on a review by the International Energy Agency (2005). The first reason is that there was no longer any staff directly assigned to the task of marketing the project. Secondly, the price was increased in that year due to a decrease in financial support for the REEP. Thirdly, the winter in 2002 was mild, so energy bills were lower and there were less indoor comfort problems. In 2003, there was a substantial rise in the number of participating households to 1130. There are also some reasons to explain this evidence: the winter of 2003 was extremely cold and long; in the fall of that year, the federal government launched their retrofit incentive; and management added another staff member who covered marketing as part of her duties. In 2004, the cumulative number of participating households exceeded 5,000, and in 2006 it reached 7,000.

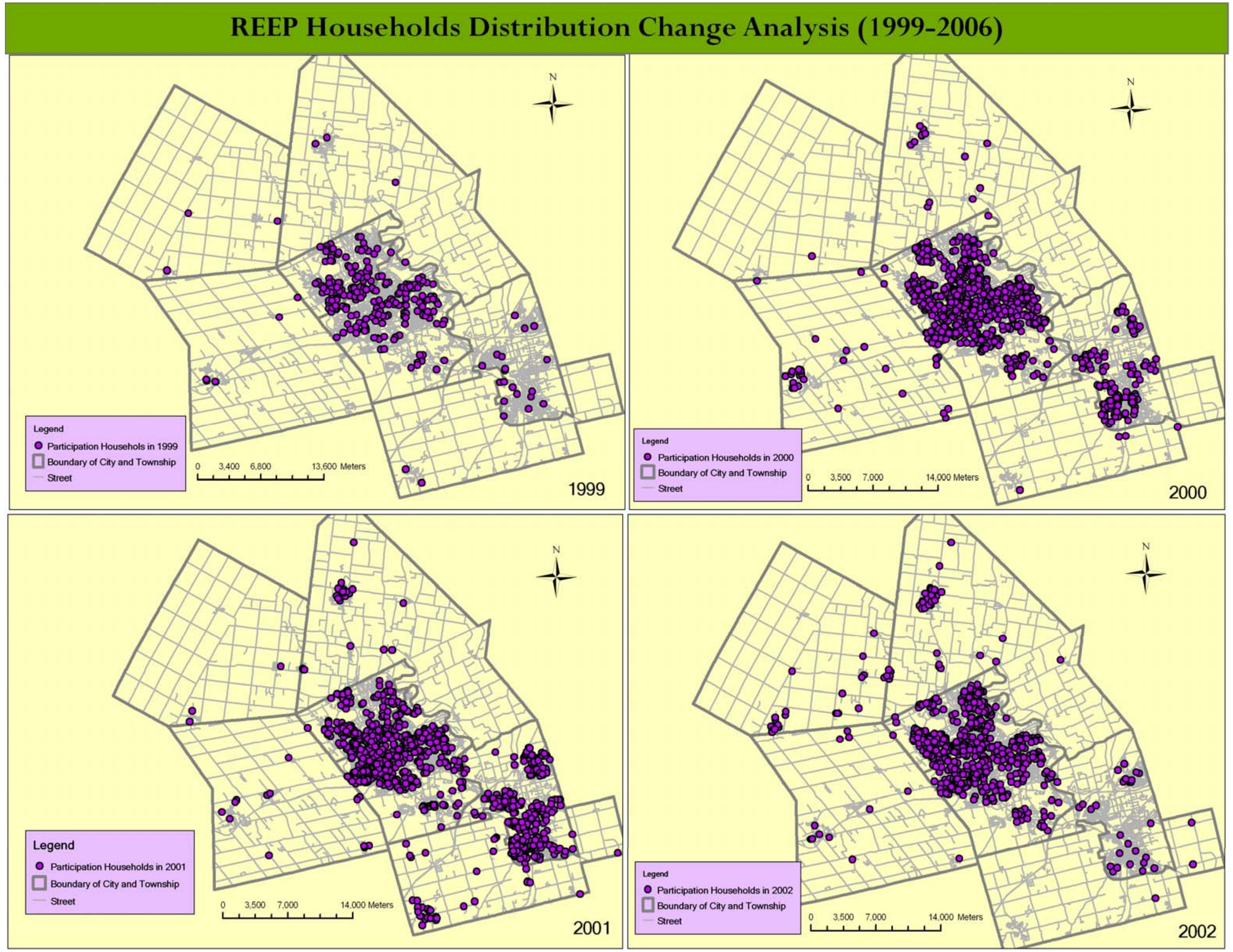


Figure 3.9 REEP Household Distribution Change Analysis (1999- 2002)

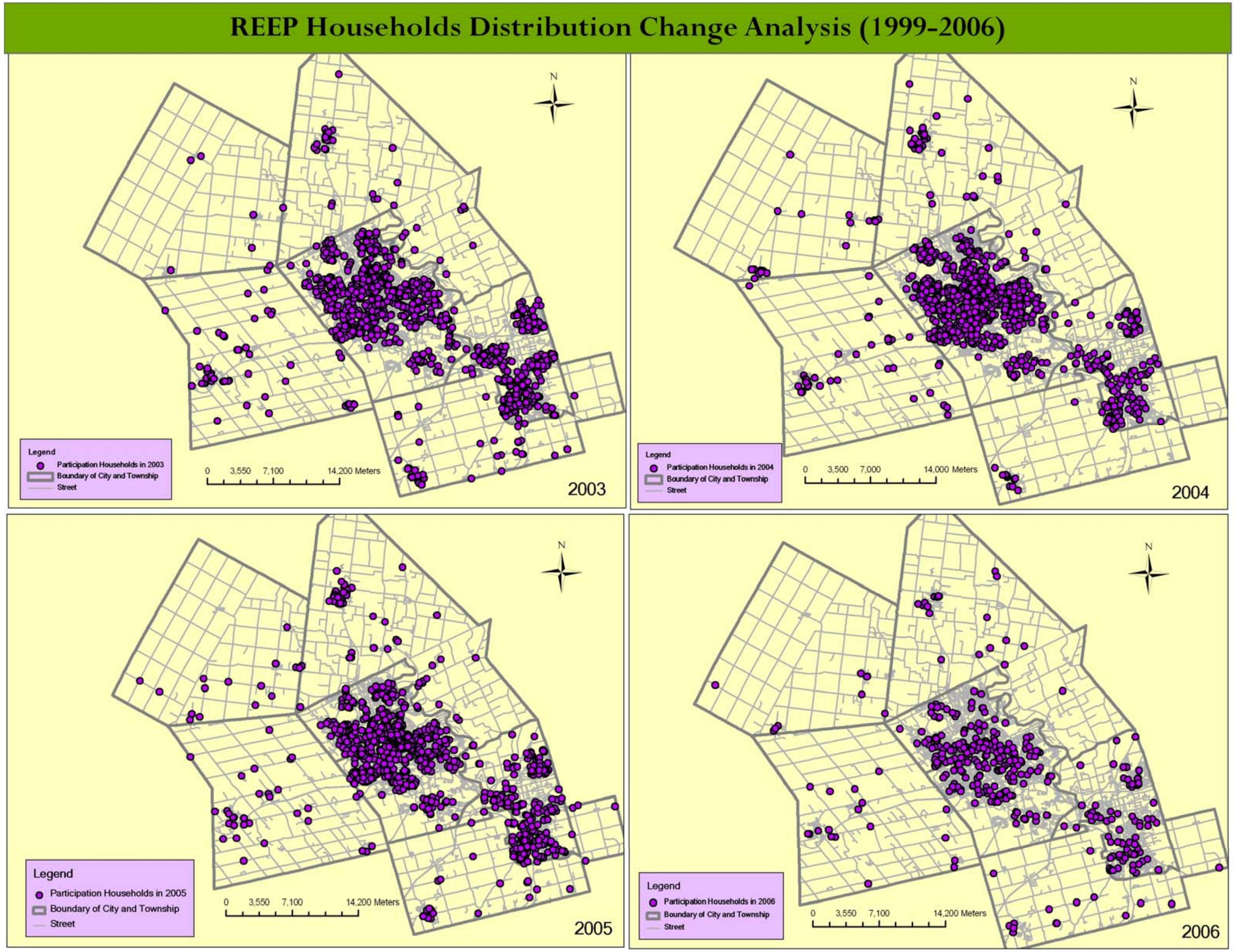


Figure 3.10 REEP Household Distribution Change Analysis (2003-2006)

3.4.2 REEP Participation Rate Analysis

Figure 3.11 to 3.14 shows participation rates at the CDA level for all eight years. Again, the pattern seems to have first developed in urban areas with a lot of high participation rate in areas near the University of Waterloo, although some areas in the City of Kitchener, which is in close proximity to the City of Waterloo, also showed relatively high participation rates.

Based on the thematic maps for participation rate, this program seems to have expanded into different areas. Initially, the participation expanded to Cambridge, and then to Wilmot (2000); by 2001 some areas in Woolwich and North Dumfries showed moderate participation rates; by 2002 this program had a lot of participating households in Wellesley. After the first four years, the program seems well developed in the whole region. From 2003 to 2006, the REEP began to gain participating households from all over the region more evenly. The majority of high participation rate locations were concentrated in the City of Kitchener and Waterloo, a few areas within Cambridge, and a few in other small communities of the township.

All participating households were asked where they heard about the REEP when they booked the evaluations for their homes. Therefore, the REEP data listed several sources of information, and REEP can track the success of each marketing initiative through these data. These sources of information include referral from friends, relatives, co-workers or neighbors; community-based events like presenting REEP information at workshops, lectures, or school visits; direct marketing through distribution of promotional materials to people's homes; community based meetings like neighborhood associations; media coverage including television, radio, and REEP website; passive marketing like front lawn sign, or REEP poster. Only a small number of households did not define how did they heard about this program.

The pie charts on the maps illustrate the percentage of participating households using each information source. Hereafter, the participation households are called evaluation, one evaluation means one participating household. In 1999, most evaluations originated from media (45.7%), followed by referral (31.3%), community-based event (8.7%), community based meeting (6.0%), direct marketing (4.9%) and passive marketing (0.4%). REEP relied heavily on media and referrals in this year. When the program first developed in the Waterloo Region, the media played a very important role in distributing information to whole region. In the first year there were around 20 print articles written, three television appearances were made and three radio pieces were broadcast about REEP (International Energy Agency, 2005). All these efforts led to the fast development of REEP. Referrals were also a key source in the first year. A significant number of households referred to REEP were from the UW community or from early partners, including staff at the Waterloo Region. In the first year of the project, REEP visited or participated in more than 50 different community based events to making direct contact with thousands of people. That makes the REEP gain 8.68% more evaluations in 1999 with passive marketing method.

In later 2000, a local utility, Kitchener Wilmot Hydro became involved in REEP and began a shift towards increased funding from utility companies in Waterloo Region. By later 2001, four local utility companies including Kitchener Wilmot Hydro, Waterloo North Hydro, Kitchener Utility and Cambridge and North Dumfries Hydro Inc had contributed over $100K to REEP (over two year period). The expansion of partner marketing of the project resulted in direct marketing being a key source from 2000 until 2003 (Appendix C shows overall comparison): more than half of the evaluations in this time period were generated by this source. This also helped to increase the number of evaluations in the whole region in year 2003 when this marketing method was developed. The project also directly assisted the utilities by offering a means to identify ways to reduce energy consumption at a time when energy bills were rising and to offer a response to high bill complaints.

As can be seen the pie charts from Figure 3.11 to 3.14, different sources of information came to play a significant role in different years. For example, after 1999 the direct marketing increased from 2000 to 2002, and decreased from 2003 due to other sources of information becoming more important. Referrals have remained as an important source of information for evaluations and increased more in percentage from 2003 to 2005. This also can explain why the REEP developed in the areas around City of Waterloo, since the project was originally developed by University of Waterloo and got referrals from people in this area.

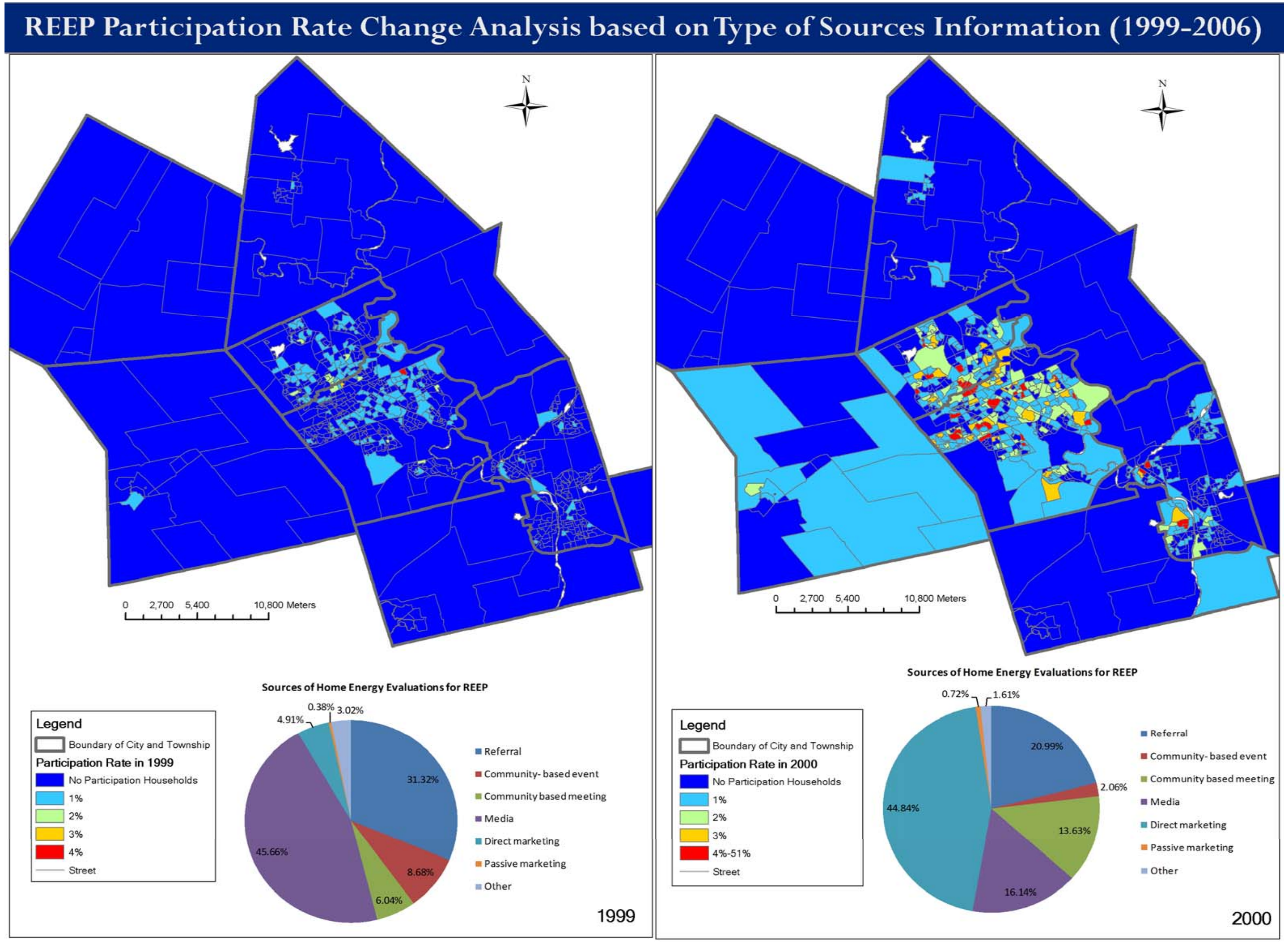


Figure 3.11 REEP Participation Rate in Each Year (1999-2000)

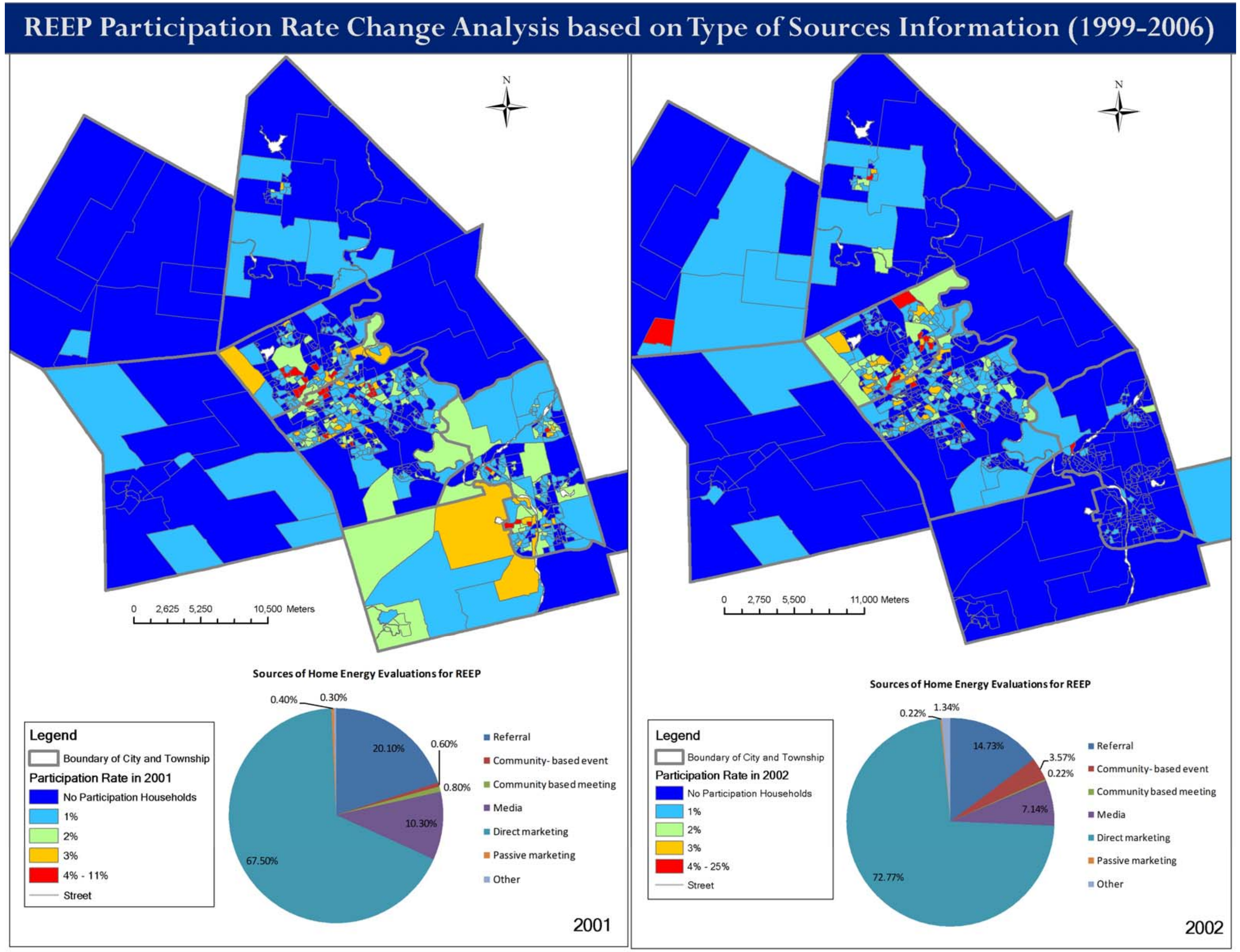


Figure 3.12 REEP Participation Rate in Each Year (2001-2002)

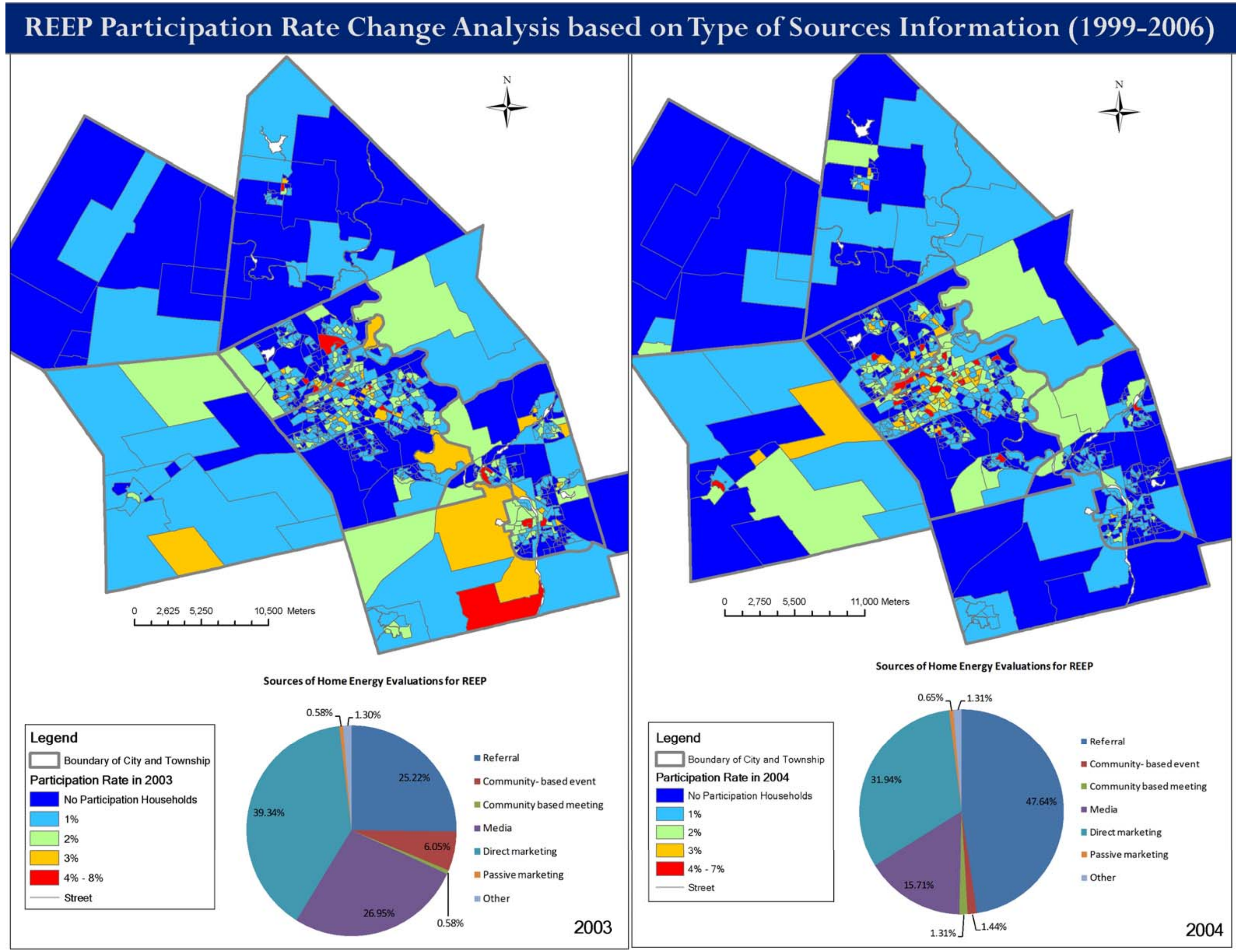


Figure 3.13 REEP Participation Rate in Each Year (2003-2004)

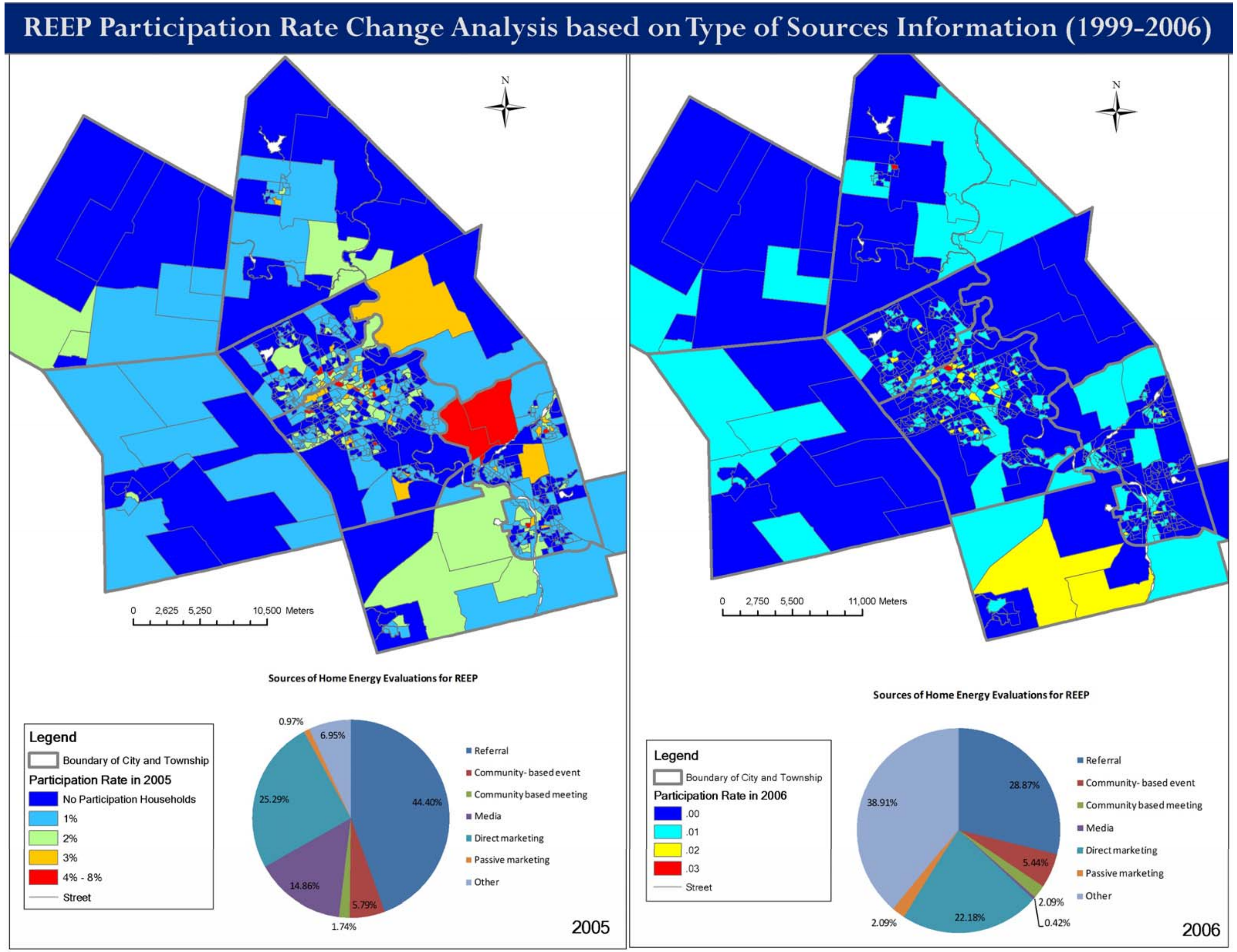


Figure 3.14 REEP Participation Rate in Each Year (2005-2006)

3.4.3 Analysis of REEP Participation Rates and Explanatory Variables

As mentioned earlier in this chapter, several factors thought to influence participation in REEP were identified from the literature review (Table 3.3). The next two chapters will focus on statistical analysis of relationships between REEP participation and the hypothesized explanatory variables. This section presents a visual interpretation of these hypothesized relationships.

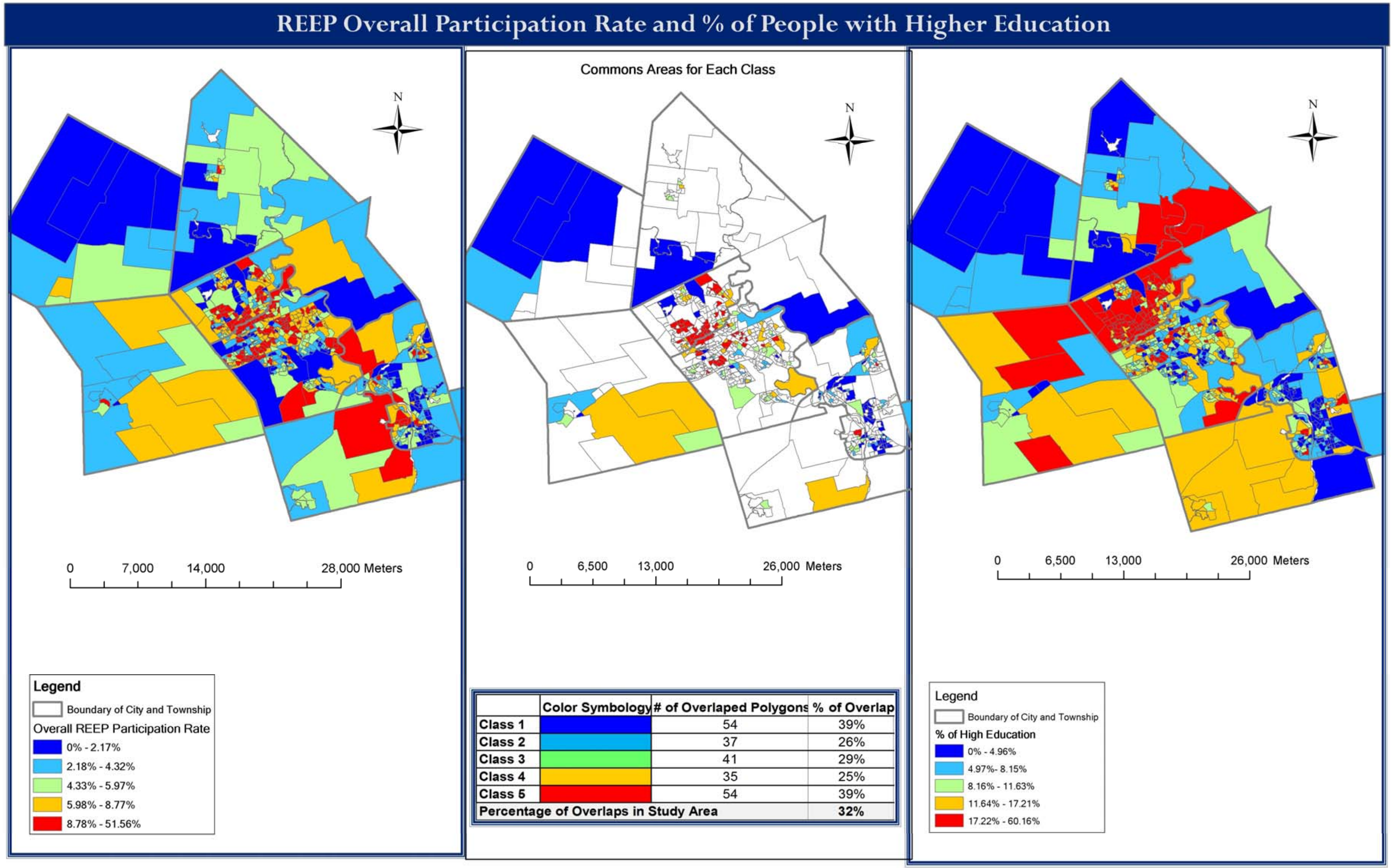
Figure 3.15 to 3.21 show the spatial distribution of each hypothesized explanatory factor in relation to the overall REEP participation rate. Each figure contains three small maps: the REEP participation rate on the left, a hypothesized explanatory factor is on the right, and the middle map shows areas of correspondence. The maps of participation rates and the hypothesized explanatory variables have been classified using quantiles and use the same color scheme. Only areas that are in the same class on both the participation rate map and the explanatory factor map are highlighted on the middle map. The summary statistics show the number and percentage of CDAs that are in corresponding classes on the participation rate and explanatory factor maps. Class 5 represents the top 20% of CDAs in terms of participation rate and the explanatory factor; class 1 represents the lowest 20% in terms of both variables.

Figure 3.15compares the spatial pattern of participation with percentage of households with higher education. High participation rates were concentrated in the City of Waterloo, north side of Kitchener and a few in the south side of Kitchener, two areas in Cambridge and the eastern side of North Dumfries, and one area in Wilmot and one in Woolwich. Low participation rates tend to occur around the fringes of the study area while moderate participation rates tend to be found in rural areas that are near the urban centers.

Areas having a high percentage of the population with higher education (at least with a university degree) covered the majority of the City of Waterloo, some areas in north Kitchener close to the boundary of Waterloo, a few in south of Kitchener, and there were also few areas covered areas in Cambridge, Wilmot and Woolwich. Areas with low percentages of the population having higher education predominate in rural areas, especially in parts of Wellesley and Woolwich.

The middle map in Figure 3.15 shows areas of correspondence for each class. Areas with high education and high participation occur in the City of Waterloo and in the City of Kitchener while areas with low education and low participation occur in rural areas, especially parts of Wellesley and Woolwich. There are areas with moderately high education and moderately high participation rates in Wilmot and scattered throughout Kitchener and Waterloo. There are also many areas that differ by only one class.

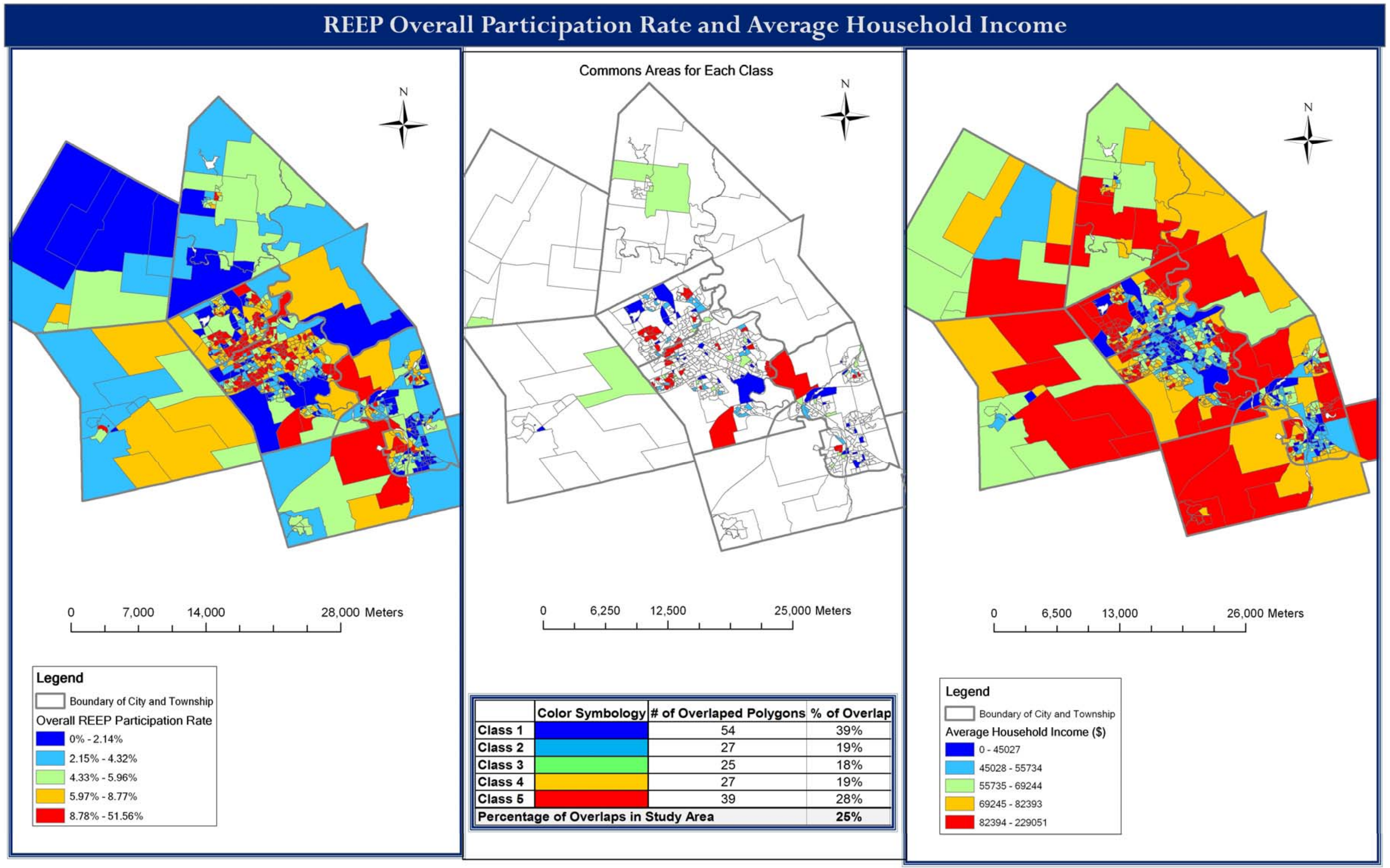
The summary statistics in Figure 3.15 show that 39% of CDAs in the top participation class are also in the top education class. Similarly 39% of CDAs in the lowest participation class are also in the lowest education class. The correspondence between intermediate classes was somewhat lower, ranging from 25% to 29%. Overall, 32% of CDAs were in corresponding classes. This indicates there may be some correlation between percentage of the population with higher education and the overall REEP participation rate.



	Color Symbology	# of Overlaped Polygons	% of Overlap
Class 1		54	39%
Class 2		37	26%
Class 3		41	29%
Class 4		35	25%
Class 5		54	39%
Percentage of Overlaps in Study Area			32%

Figure 3.15 Compare % of Higher Education with REEP Overall Participation Rate

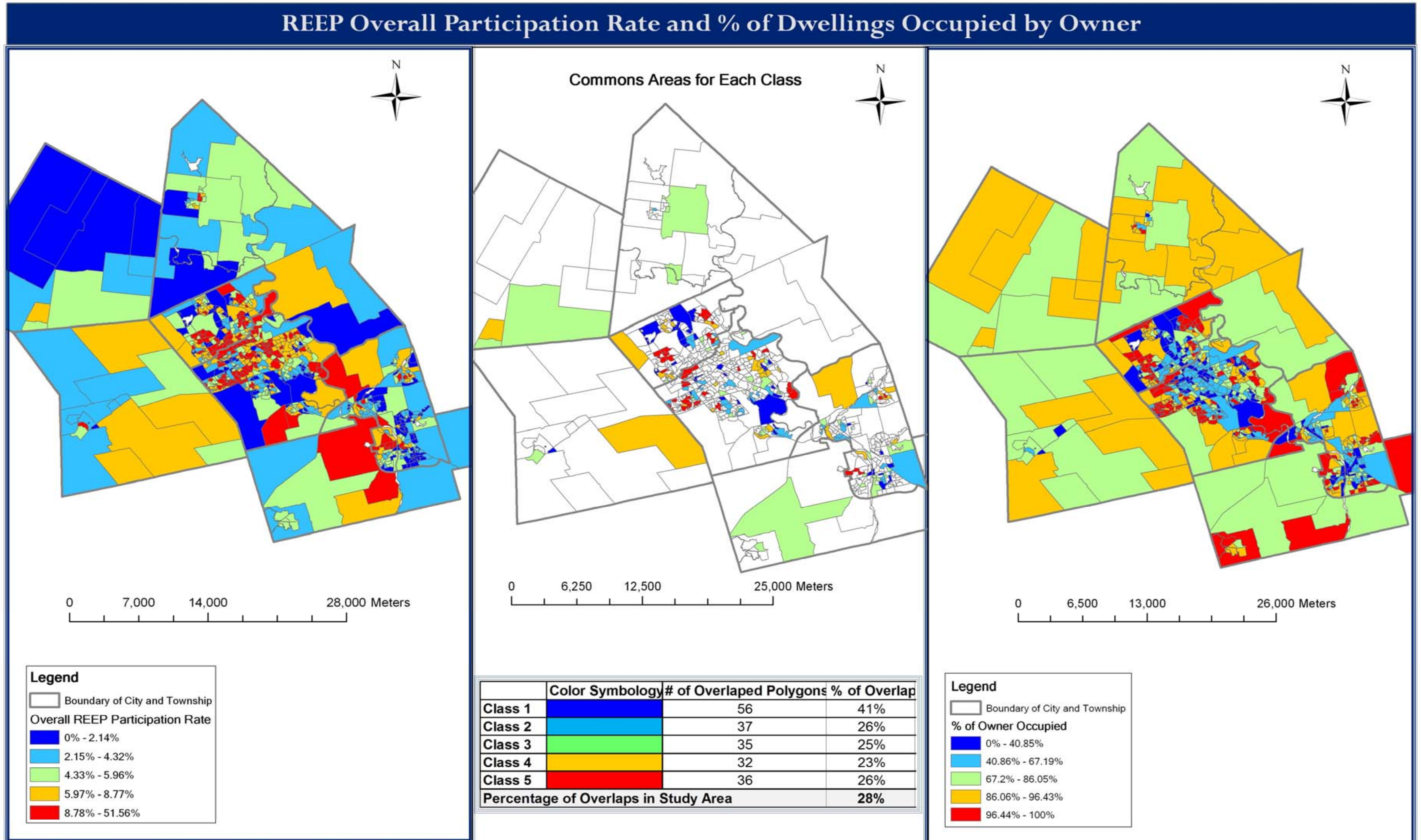
Figure 3.16 shows the relationship between participation rate and average household income. High income areas surround the three urban centers. This may reflect the preference of some high income households for an ex-urban lifestyle. These areas tend to correspond to areas having moderate to moderately high participation rates. Thus there is not an exact correspondence. Areas where there does appear to be correspondence between the two maps include high income/high participation CDAs within the urban areas, low income/low participation CDAs within the urban areas, and a few rural CDAs with moderate incomes and participation. This suggests a weaker relationship than was the case with participation and education. This is confirmed by the summary statistics which indicate that only 25% of CDAs are in the same class on both maps. Based on visual and statistical comparison, this information indicates there may not be a strong correlation between average household income and the overall REEP participation rate.



	Color Symbology	# of Overlaped Polygons	% of Overlap
Class 1		54	39%
Class 2		27	19%
Class 3		25	18%
Class 4		27	19%
Class 5		39	28%
Percentage of Overlaps in Study Area			25%

Figure 3.16 Compare Average Household Income with REEP Overall Participation Rate

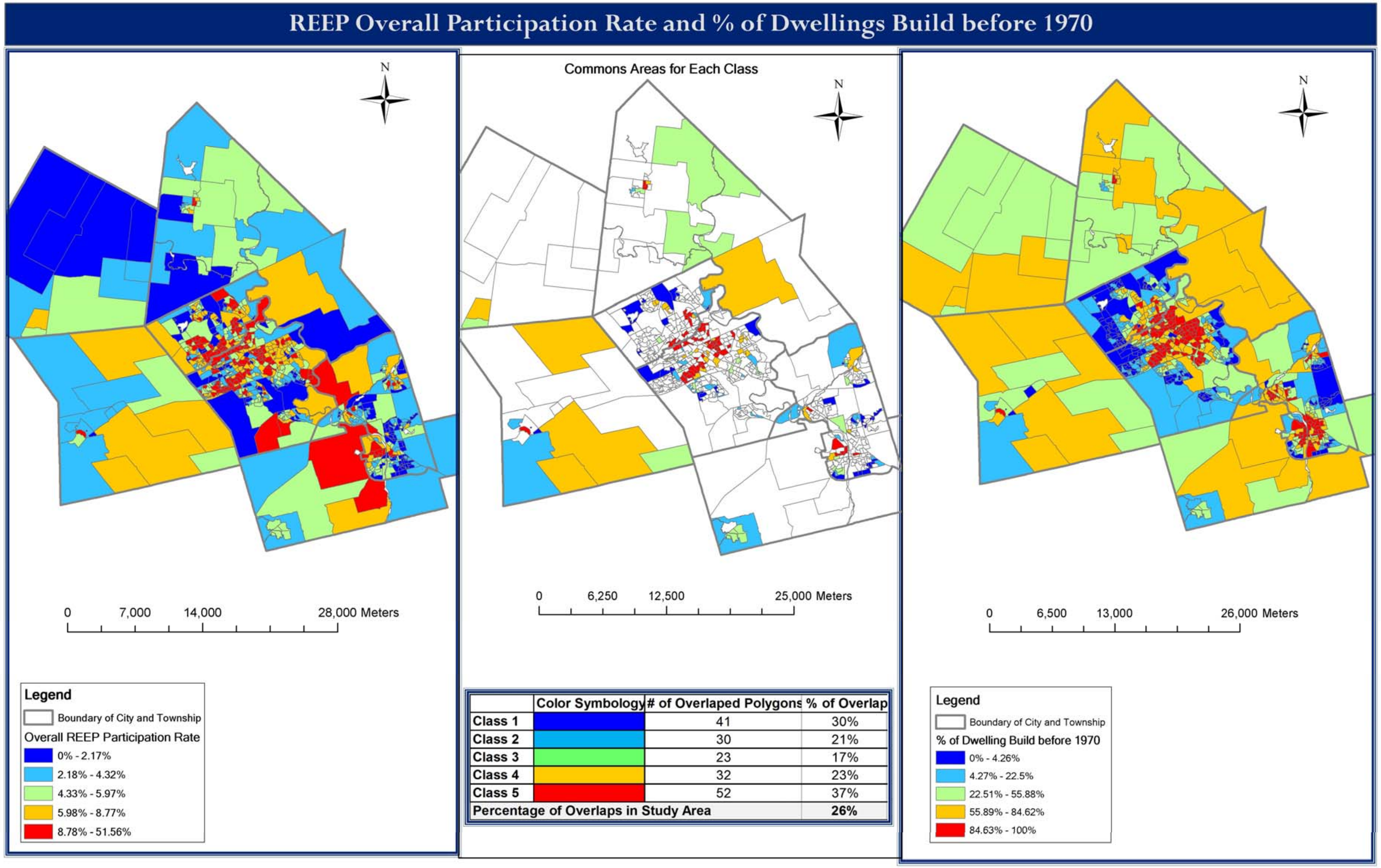
Figure 3.17 compares participation rates with percentage of owner-occupied dwellings. The percentage of owner-occupied dwellings tends to be moderate to high in the rural areas, but is low to moderately low in much of the urban areas reflecting the higher percentage of households living in rented accommodation. The middle map shows areas of correspondence in all five classes within the urban areas while in the rural areas there are several CDAs that are in the moderate or moderately high class on both maps. The summary statistics indicate that the strongest correspondence is in the low participation – low ownership category: 41% of low participation CDAs are also low ownership CDAs. In the other categories, 23% to 26% of CDAs were in the same category on both maps. These results suggest a moderate association, stronger than for household income but weaker than for education.



	Color Symbology	# of Overlaped Polygons	% of Overlap
Class 1		56	41%
Class 2		37	26%
Class 3		35	25%
Class 4		32	23%
Class 5		36	26%
Percentage of Overlaps in Study Area			**28%**

Figure 3.17 Compare % of Owner-Occupied with REEP Overall Participation Rate

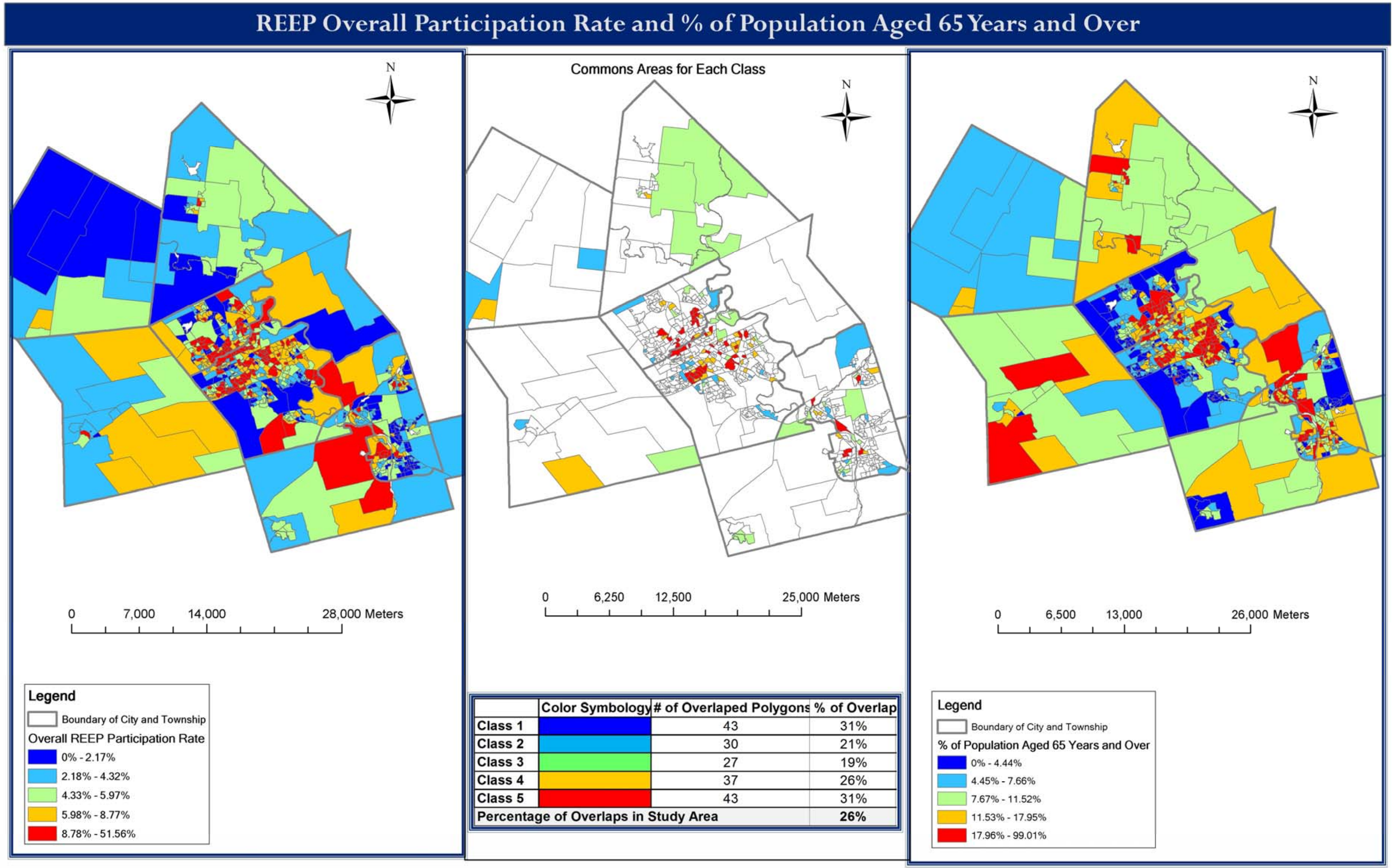
Figure 3.18 compares participation rates with percentage of dwellings built before 1970. The urban core areas have high percentages of older buildings, rural areas are mostly in the moderate to moderately high categories, and suburban areas within Kitchener, Waterloo and Cambridge generally have low percentages of buildings built before 1970. Based on the summary statistics, the strongest correspondence occurs in the high participation – high percentage of older buildings class that includes may urban core CDAs in Kitchener and Waterloo, plus a few outliers in Cambridge, Elmira and New Hamburg. There are also several CDAs in the low participation – low percentage of older buildings category. These tend to be located around the edges of the urban areas and appear to be areas that may be industrial land or are areas that are not yet fully developed. Again these results suggest a moderate association between age of dwellings and participation in REEP.



	Color Symbology	# of Overlaped Polygons	% of Overlap
Class 1		41	30%
Class 2		30	21%
Class 3		23	17%
Class 4		32	23%
Class 5		52	37%
Percentage of Overlaps in Study Area			**26%**

Figure 3.18 Compare % of Dwellings Built before 1970 with REEP Overall Participation Rate

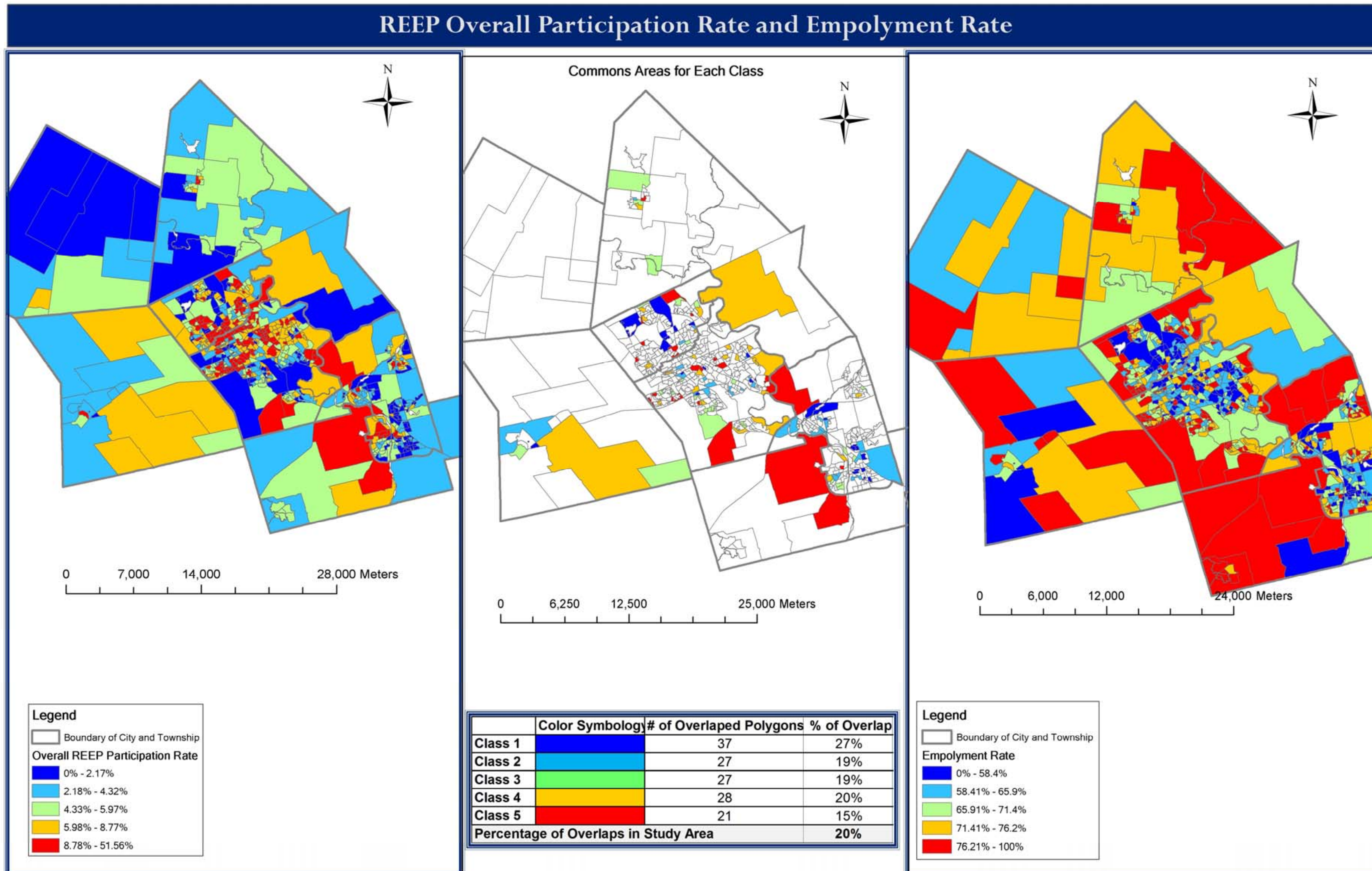
Figure 3.19 compares participation rates with percentage of the population aged 65 years and over. The urban core areas have high percentages of population aged 65 years and over, plus a few outliers in Woolwich and Wilmot. Rural areas are mostly in the moderate to moderately low categories, and suburban areas within Kitchener, Waterloo and Cambridge generally have moderately low percentages of population aged 65 years and over. The middle map shows areas of correspondence in all five classes within the urban areas there are some CDAs that are in the high or moderately high class on both maps. The summary statistics in Figure 3.19 show that 31% of CDAs in the top participation class are also in the top aged population class. Similarly 31% of CDAs in the lowest participation class are also in the lowest aged population class. The correspondence between intermediate classes was somewhat lower, ranging from 19% to 26%. Overall, 26% of CDAs were in corresponding classes. These results suggest a moderate association, stronger than for household income but weaker than for education, age of dwellings and ownership.



	Color Symbology	# of Overlaped Polygons	% of Overlap
Class 1		43	31%
Class 2		30	21%
Class 3		27	19%
Class 4		37	26%
Class 5		43	31%
Percentage of Overlaps in Study Area			26%

Figure 3.19 Compare % of Population Aged 65 Years and Over with REEP Overall Participation Rate

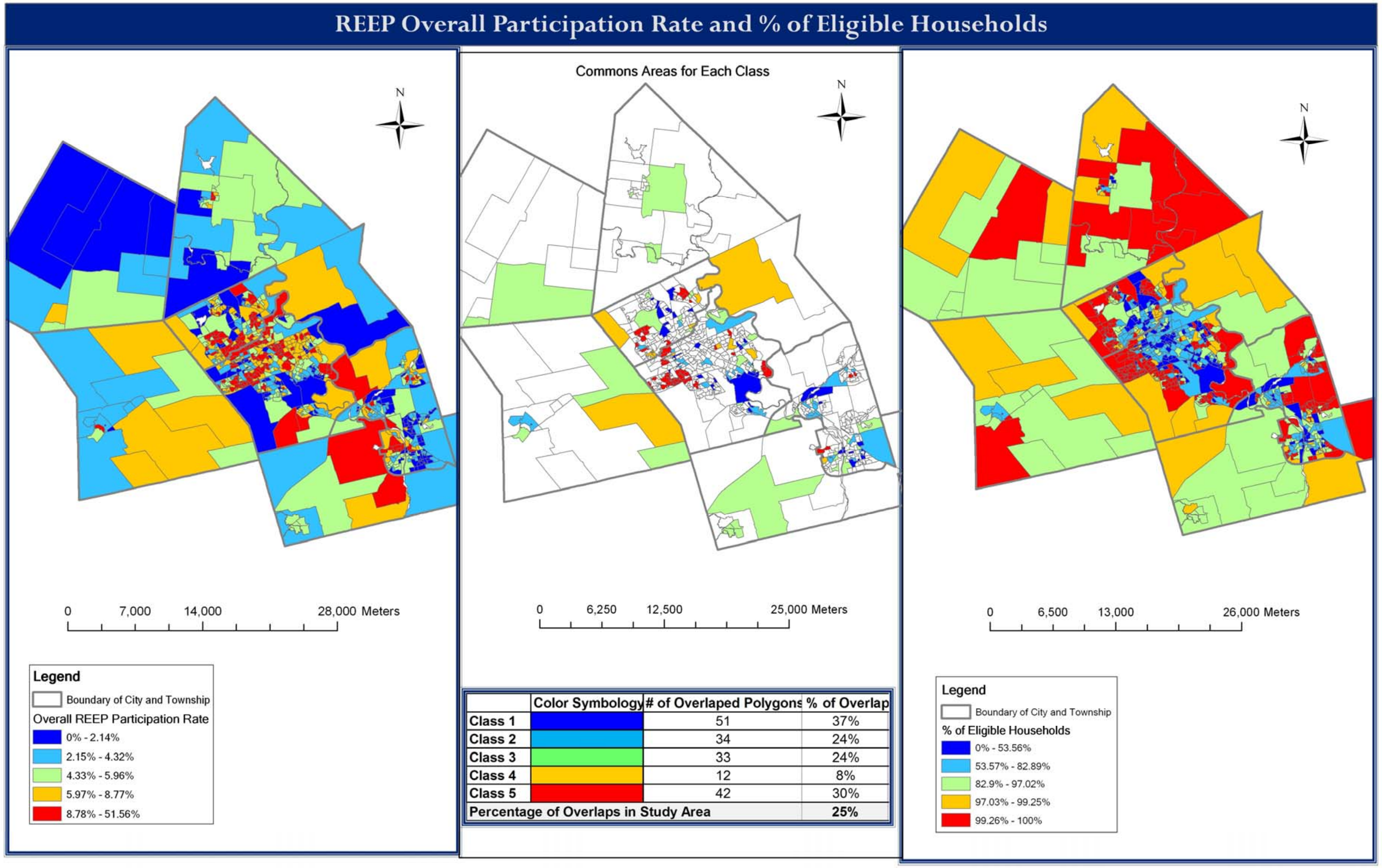
Figure 3.20 compares participation rates with employment rates. The urban core areas have some CDAs with high employment rates, rural areas are mostly in the moderately high to high categories, especially in North Dumfries almost cover the entire areas, and suburban areas within Kitchener, Waterloo and Cambridge generally have high employment rates as well. While, low to moderately low employment rates tend to occur in the three urban centers, plus a few in Wellesley, Wilmot, Woolwich and North Dumfries. The summary statistics indicate that a relative strong correspondence is in the low participation – low employment rate: 27% of low participation CDAs are also low employment rate CDAs. In the other categories, 15% to 20% of CDAs were in the same category on both maps. These results suggest a very weak association.



	Color Symbology	# of Overlaped Polygons	% of Overlap
Class 1		37	27%
Class 2		27	19%
Class 3		27	19%
Class 4		28	20%
Class 5		21	15%
Percentage of Overlaps in Study Area			20%

Figure 3.20 Compare Employment Rate with REEP Overall Participation Rate

Figure 3.21 compares participation rates with percentage of eligible households. The moderately high to high percentage of eligible households occurs in rural areas and also the suburban areas of Kitchener, Waterloo and Cambridge, and a few in the three urban centers as well. Whereas, low to moderately low percentage of eligible households tend to occurs in the three urban centers, plus a few outliers in Wilmot and Woolwich. The summary statistics indicate that the strongest correspondence is in the low participation – low percentage of eligible households: 37% of low participation CDAs are also low percentage of eligible households CDAs, while, 30% of high participation CDAs are also with high percentage of eligible households CDAs. The overall area of correspondence is 25%. These results suggest a moderate association, stronger than employment rate but weaker than other variables.



	Color Symbology	# of Overlaped Polygons	% of Overlap
Class 1		51	37%
Class 2		34	24%
Class 3		33	24%
Class 4		12	8%
Class 5		42	30%
Percentage of Overlaps in Study Area			25%

Figure 3.21 Compare % of Eligible Households with REEP Overall Participation Rate

The last line in the statistics tables presented in the above maps provides a crude index of the degree of association between participation rates and the hypothesized explanatory variables. The reported percentages are based on the number of CDAs in each quantile on the participation rate map that are in the same quantile on the explanatory factor map. Thus, higher percentages indicate stronger associations. Comparing these statistics indicates that the percentage of population with higher education has the strongest association (32%), followed by the percentage of owner-occupied dwellings (28%), percentage of dwellings build before 1970 and percentage of population aged 65 years and over (26%), average household income and percentage of eligible households (25%), and the employment rate (20%). These results indicate that all of the hypothesized explanatory variables have an impact on participation rates, although the effect of employment rates is weaker that the other factors.

3.4.3.1 Correlation Analysis

To confirm these results, Spearman's rank correlation coefficients (r) were calculated between number of participant households and each explanatory variable (Table 3.4). The *r* for the number of people with higher education was 0.454, meaning that there is a moderate positive relation between higher education and participation in REEP, CDAs with a larger number of households with higher education are expected to have higher participation rates. Similarly, the *r* values for number of owner-occupied dwellings (0.453), average household income (0.417), number of eligible households (0.405), indicate moderate positive relationships. Weaker positive relationships were found for population aged 65 and over (0.238), dwellings built before 1970 (0.186), and number of people employed (0.093). All results are significant at the 0.01 level except for number of people employed which is significant at the 0.05 level but not at the 0.01 level.

Table 3.4 Correlation With Participation Rate

Name of Factors	Test Factors	Spearman's rho Correlation Coefficient
Population with Higher Education	**Bache_or_H**	0.454**
Average Household Income	**Avhhinc**	0.417**
eligible households (dwellings)	**Appr_Dwe**	0.405**
Employment Rate	**Employ_Rte**	0.093*
Dwellings Occupied by Owner	**Owned**	0.453**
Population Aged 65 Years and Over	**Over65**	0.238**
Dwellings Built Before 1970	**Bf1970**	0.186**

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

The above correlation analysis generally confirms the results obtained based on map interpretation and indicates that, with the possible exception of employment rate, all hypothesized explanatory variables show a strong enough degree of association with participation rates to warrant inclusion in the regression analyses reported in chapters 4 and 5.

3.5 Summary

This chapter has presented a visual interpretation of the spatial patterns of REEP participation in Waterloo Region, the impact of sources of information on participation, and relationships between participation rates and factors identified from the literature review that may influence participation rates for this program.

After introducing the study area and reviewing the data sources available for this research, the methods used to convert the data into a GIS database suitable for performing spatial analysis were described. These included geocoding the REEP households, spatial joining the REEP households to the census dissemination areas to associate a CDA identifier with each households, calculating participation rates for each CDA, and converting selected demographic variables at the CDA level from counts to percentages.

The analysis in this chapter focused on interpretation of three sets of thematic maps designed to address three research hypotheses:

1. That the spatial pattern of participation changed over time and tended to cluster in urban areas;

2. That the spatial distribution differs by source of information; and

3. That the relationships exist between participation rates and explanatory variables.

Firstly, visual comparison was used to evaluate whether the spatial pattern of participation changes over time. Generally speaking, the participating households tend to be clustered in urban areas, and other small communities like Elmira, Wellesley, Baden, New Hamburg, New Dundee, and Ayr; and tend to be scattered in rural areas. In addition, there was a trend for the changes in the spatial distribution of participation. Participating households first developed from the City of Waterloo and City of Kitchener, and then expanded to the City of Cambridge, to form a very dense pattern through eight years.

Secondly, the pie chart and thematic maps for participation rates based on each year provided insight into the changes of the source of information from 1999 to 2006. According to the visual comparison, different sources of information came to play a significant role in different years; also the spatial distribution differs by source of information. For example, after 1999 the direct marketing increased from 2000 to 2002, and decreased from 2003 as other sources of information became more important. Referrals have remained as an important source of information for evaluations and increased more in percentage from 2003 to 2005.

Finally, Spearman’s rank correlation was used to assess the strength of associations between hypothesized explanatory factors and participation rates. In general, these results confirmed the findings based on map interpretation and indicate moderate positive associations between participation rates and each of the explanatory factors. The following chapters will present the results of multivariate analysis of the factors affecting participation in the REEP.

Chapter 4 Poisson Regression Analysis

4.1 Introduction

This chapter applied Poisson Regression to analyze the relationship between participating REEP households and the hypothesized explanatory variables. First of all, this chapter discusses how this method works and why the method is appropriate for the data used in this thesis. The second section presents the results of the analysis and is followed by discussion and interpretation of these findings. Finally, the chapter concludes with a discussion of the limitations of the method and a short summary.

4.2 Methodology

Regression analysis constructs an explanatory or predictive model of a dependent or response variable on the basis of one or more independent or explanatory variables. Poisson regression is appropriate when the dependent variable is a count, such as the number of times an event occurs or the number of people in a certain category (Cameron & Trivedi, 1998; Strien, Pannekoek, Hagemeijer & Verstrael, 2004).

In this research, the response variable was count data, and the count data were not normally distributed (see Appendix D). Since the dependent variable is in the form of a count, the values are all positive integers, the range of values is limited since counts cannot be less than 0, variances increase as the number of eligible household increases, and the data are skewed rather than normal. Therefore standard linear regression is inappropriate and transformation of counts is often unsatisfactory, especially in case of many zero counts (Strien *et al.*, 2004). For event or count data, the Poisson distribution better describes the data distribution and Poisson Regression provides a more appropriate analysis model.

The above explanations show why the Poisson regression has been selected for this part of the analysis. The SPSS software (version 16.0, SPSS Inc.) was utilized to conduct this analysis. In this software, it treats the dependent variable with an equal weight; however, the numbers of eligible dwellings in each CDA were not equal. The Weight Case function was applied, the weight variable created from the number of eligible dwellings in each CDA divided by the mean of number of eligible dwellings. This function weighted the dependent variable to adjust for the number of potential REEP participants in each CDA.

The Poisson regression model is sometimes known as a log-linear model which is a form of Generalized Linear Model (GLM) (Cameron and Trivedi, 1998; McCullagh and Nelder, 1989; Dobson, 1991). Thus, if there are several independent variables X_1, $X_2 \ldots X_n$ and we let X denote the vector s(X_1, $X_2 \ldots X_n$), the predicted value of the dependent variable Y for case i is given by the equation

$$\text{Log}_e(y\hat{}_i) = \beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}\ldots+ \beta_n X_{ni} \text{ (eq 4.1)}$$

Where $y\hat{}_i$ is the predicted value of Y for case i. $X_i= (X_{1i}, X_{2i}\ldots X_{ni})$ is the realization of X for case i. β_0 is the intercept- the value of y^ when X is a zero vector, and β_1, $\beta_2\ldots$ are the weights of X_1, X_2 …the amount by which log(y^) changes as X increases by one unit.

The independent variables at the CDA level were used in this analysis (Table 4.1). The response variable is the total number of participating households in each CDA during the entire study time period (**T_HH_IN_Pa**). There are a total of 694 CDAs in this study. Poisson Regression analysis was used to explore the relationship between number of participating households and percentage of population with higher education (**per_high_e**), average household income (**avhhinc**), employment rate (**employ_rte**), percentage of Dwellings Occupied by Owner (**owned_perc**), percentage of Population Aged 65 Years and Over (**perover65**), percentage of Dwellings Built Before 1970 (**perbf1970**), and/or number of eligible households (**appr_dweno**) in each CDA.

Percentages were used for some independent variables to simplify interpretation and discussion of the results.

Table 4.1 List of Independent (predictor) Variables

1. % of Population with Higher Education
2. Average Household Income
3. Employment Rate
4. % of Dwellings Occupied by Owner
5. % of Population Aged 65 Years and Over
6. % of Dwellings Built Before 1970 (before energy crisis)
7. Number of eligible households (dwellings)

4.3 Results and Discussion

Table 4.2 provides information about the specified model and the input data set. The response distribution is specified as Poisson, and the link function is chosen to be log. Usually, in the Poisson regression, the scale parameter is assumed to be 1. McCullagh and Nelder (1989) use the Pearson chi-square estimate to obtain more conservative variance estimates and significance levels. Therefore, in this case, the Pearson chi-square was selected as the method for estimating the scale parameter.

Table 4.2 Model Information

Dependent Variable	T_HH_IN_PA
Probability Distribution	Poisson
Link Function	Log

Figure 4.1 and 4.2 can be used to diagnose how well the data fit the Poisson regression model. Usually, no pattern in the residual scatter plot indicates that the data possibly fit this model well. Comparing the two scatter plots obtained for the weighted and unweighted cases, Figure 4.1 seems more random than Figure 4.2. Thus, the weighted dependent variable seems to improve the quality of data fit for Poisson regression model.

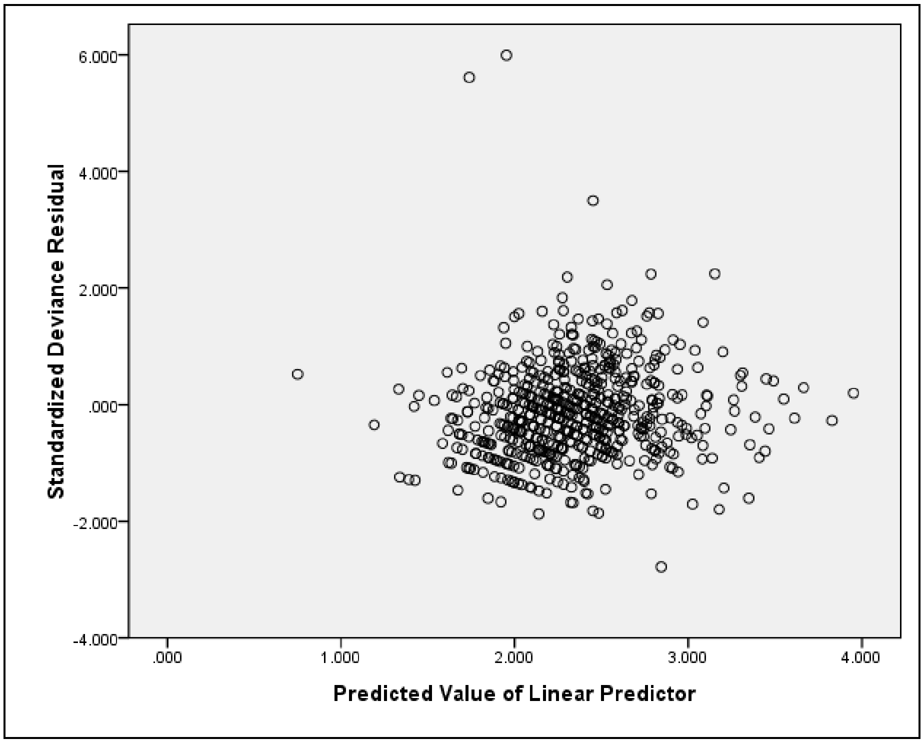


Figure 4.1 A residual Diagnostic Plot (weighted case)

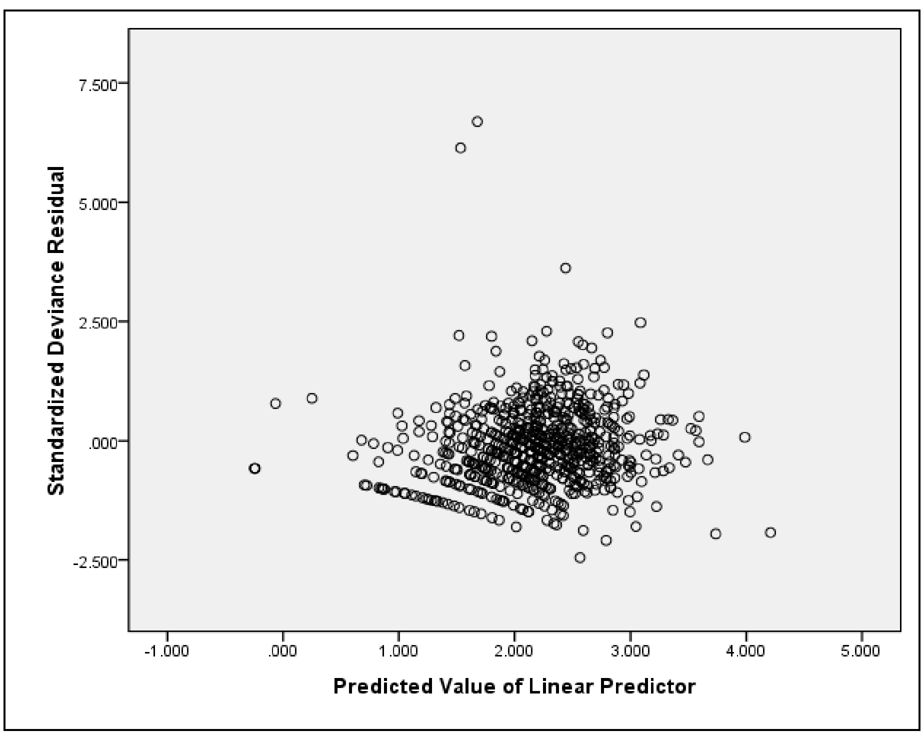


Figure 4.2 A residual Diagnostic Plot (unweighted case)

Table 4.3 shows the Analysis of Parameter Estimates, which summarize the results of the parameter estimation process. For each parameter in the model, SPSS displays columns with the parameter's name, the degrees of freedom associated with the parameter, the estimated parameter value, the standard error of the parameter estimate, the confidence intervals, and the Wald chi-square statistic and associated *p*-value for testing the significance of the parameter to the model.

The **B** column in Table 4.3 displays the estimated Poisson regression coefficients for the model. Recall that the response variable is a count variable (in this case, the total number of participating households in each CDA), and Poisson regression models the log of the expected count as a function of the predictor variables.

The **Sig.** column in Table 4.3 provides the *p*-values of the coefficients or the probability that, within a Poisson regression model, the null hypothesis that a particular predictor's regression coefficient is zero given that the rest of the predictors are in the model. They are based on the Wald Chi-Square test statistics of the predictors.

Table 4.3 Analysis of Parameter Estimates (Weighted dependent variable by weight_M)

Parameter	**B**	Std. Error	Hypothesis Test		
			Wald Chi-Square	df	**Sig.**
(Intercept)	**.263**	.2727	.933	1	**.334**
AVHHINC	**3.188E-6**	1.2402E-6	6.609	1	**.010**
APPR_DWENO	**.001**	9.8424E-5	118.430	1	**.000**
PER_HIGH_E	**.029**	.0030	98.230	1	**.000**
EMPLOY_RTE	**.008**	.0036	5.662	1	**.017**
OWNED_PERC	**.005**	.0015	11.927	1	**.001**
PEROVER65	**.009**	.0042	4.428	1	**.035**
PERBF1970	**.006**	.0009	41.700	1	**.000**
(Scale)	4.980[a]				

Dependent Variable: T_HH_IN_PA * Weighted by Weight_M (Weight_M= number of eligible households in each CDA/mean of eligible households in study area)

Model: (Intercept), PER_HIGH_E, AVHHINC, EMPLOY_RTE, OWNED_PERC, PEROVER65, PERBF1970, PER_ELGIHO

a. Computed based on the Pearson chi-square.

4.3.1 Results

Based on Table 4.3, all the variables were significant in terms of predicting the number of participation households in each CDA. The Poisson regression coefficients (B) are estimates of the expected increase in the log of the number of participants for a unit increase in the value of the independent variable. For example, a one percent increase in the percentage of people having higher education is expected to increase the log of the number of households participating in REEP by 0.29, assuming the other variables in the model are held constant. Thus we would expect a CDA with a larger percentage of the population having higher education to also have a higher participation rate. Since all coefficients are positive, similar interpretations can be made for the other variables, although the expected increases in the log of the number of participating households would be smaller because the other coefficients are smaller.

The Wald Chi-Square test statistic testing the slope for PER_HIGH_E on T_HH_IN_Pa (the total number of participation household in each CDA) is zero, given the other variables are in the model, is $(0.029410/ 0.002967)^2 = 98.230$, with an associated *p*-value of less than 0.0001. The 95% confidence interval was set in this model; this result would reject the null hypothesis and conclude the Poisson regression coefficient for PER_HIGH_E is very statistically different from zero given other variable in the model.

The p-values indicate that all predictor variables are significant at the 95% confidence level. Thus, in each case, we would reject the null hypothesis of no association between the independent and dependent variables when the other variables in the model are held constant.

To assess the predictive ability of the above Poisson regression model, the model was used to calculate the expected number of REEP participants in each CDA given observed values for the independent variables. Simple correlation analysis was used to estimate the strength of the relationship between the predicted number of participants and the actual number of participants. The Spearman's rho correlation coefficient of 0.669 indicates that there is a relatively strong positive relationship between predicted and actual numbers (see Table 4.4). Furthermore, the R squared value of 0.44756 means that 44.8% of the variation in number of participating households can be explained by the Poisson regression model. Therefore, all the independent variables are useful for predicting the household participation in the REEP.

Table 4.4 Correlations Test

			T_HH_IN_PA	Predict_P_HH
Spearman's rho	T_HH_IN_PA	Correlation Coefficient	1.000	.669**
		Sig. (2-tailed)	.	.000
		N	660	660
	Predic	Correlation Coefficient	.669**	1.000
		Sig. (2-tailed)	.000	.
		N	660	660

**. Correlation is significant at the 0.01 level (2-tailed).
*R square = $(0.669)^2 = 0.44756$

4.3.2 Overall Discussion

In the socio-demographic factors, both the percentage of people with higher education and percentage of dwellings occupied by owners have *p*-value of less than 0.001. That indicates these variables are highly significant. The other variables such as average household income, employment rate, and percentage of population aged 65 years and over have *p*-values of 0.010, 0.017, and 0.035 respectively, which also indicate that these variables are significant, but not as significant as the first two.

In the physical factors, both the percentage of dwellings built before 1970 and the total number of eligible dwellings have *p*-values less than 0.001that indicates that these variables are highly significant. It is interesting to note that all seven factors of this study show significant associations with participation in the REEP when all predictor variables are included in the same Poisson regression model.

A possible explanation for the relationship between higher education and participation in the REEP could be that people with higher education may better understand the goals of REEP and the benefits that participation in the REEP will bring to them. They also may be more willing to learn or accept new things, and participate in programs like REEP. Some of them may have learned about the energy efficiency concept during school years, and others may know this concept through research or different sources. Higher education people may still have contact with university or other university's classmates that they also can get information about REEP from referral. This was one of the most important sources in REEP marketing development.

There is also a logical explanation for the association between home ownership and participation in the REEP. People who own their dwelling may live in it for a long time while people who rent a dwelling tend to move more frequently. Participating in the REEP may require a large investment at the beginning that includes the cost of the home evaluation and any needed upgrades and which may have a relatively long payback period. The owner also gains resale value of the home. Therefore, only people who live in the dwelling for a long time may be interested in this project and people who rent the dwelling may not care about the energy performance in the dwelling.

Household income was also significantly associated with participation in the REEP. There are three possible explanations for this relationship. First, the REEP home evaluation requires payment of a service charge which low income households may be unwilling to pay. Furthermore, the cost of upgrades may be prohibitively expensive, especially for the homes with very low energy efficiency. Finally, households may have different desires regarding their standard of living. High income families who would like to achieve high standard of living may care more about indoor comfort and be more willing to use advanced energy systems for their home.

The current study also found that employment rate is significant in determining participation in the REEP. If the employment rate increases that means there are more people in the labor force. In chapter 3, the source of information of REEP has been discussed, and referral was one of the most important information sources. Therefore, if more people are in the labor force, they may absorb more information about the society, including REEP information. Once they become familiar with this project, the benefits from joining this project may attract them, and they may decide to participate.

There are three reasonable explanations for the relationship between populations aged over 65 years and over and participation in the REEP. People who live longer may be more aware of environmental changes (climate change) and may want to preserve the environment for later generations. For example, 15 years ago they may not have used air conditioners, but now use air conditioners every summer. They can feel the climate changes. The second important reason may be because population aged 65 years and over have more extra time during the day. Since the energy efficiency project requires time to evaluate their homes, if they have more flexible time they may be more likely to participate in this project. The last possible reason is they may stay home longer than other people, thus indoor comfort problems become more serious for them, and also by joining the REEP they can save more money since they may require higher temperature in their home during winter.

The percentage of dwellings built before 1970 is also positively related to participation in the REEP in the Passion regression model. Usually, older houses have lower energy efficiency. The house may have a lot of problems such as air leakage, low efficiency furnace, poor wall insulation and so on. These older houses may have much more space to upgrade than newer houses. These houses also have high potential for energy and money savings by doing upgrades. Furthermore, some older houses may already require changes in energy system in the house, since the systems are old. The houses may also have more problems to deal with concerning indoor comfort.

The explanation for the relationship between total number of eligible dwellings in each CDA and participation in the REEP is pretty obvious: higher participation is to be expected in the areas that have a larger pool of eligible dwellings.

4.4 Limitation

The Poisson distribution has three special features: it is skewed, non-negative, and has a variance equal to the mean (Cameron & Trivedi, 1998; McCullagh & Nelder, 1989). The last characteristic is the main assumption for Poisson regression: if the mean and variance are not equal, then the over-dispersion may apply for the data (Berk & MacDonald, 2007). The descriptive statistics presented in Table 4.5 suggest that this may be the case for the data used in this analysis.

Table 4.5 Descriptive Statistics –Number of Participating Households Based on CDA Level

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
T_HH_IN_PA	694	0	66	12.33	9.462	89.538

Over-dispersion is a phenomenon that sometimes occurs in data that are modeled with Poisson distributions (Cameron and Trivedi, 1998). In SPSS, information on the goodness of fit can be used to discover the potential for fitting an over-dispersion Poisson regression model.

Table 4.6 displays the goodness of fit statistics that summarize the fit of the Poisson regression. For a Poisson regression the Value/df for Deviance and Pearson Chi-Square statistics should be near 1.0. If the Value/df is greater than 1, then the data are over-dispersed. The Value/df for Deviance and Pearson Chi-Square were 3.725 and 4.980 respectively, both indicating over-dispersion.

Table 4.6 Goodness of Fit

	Value	df	Value/df
Deviance	2682.147	720	3.725
Scaled Deviance	538.590	720	
Pearson Chi-Square	3585.560	720	4.980
Scaled Pearson Chi-Square	720.000	720	
Log Likelihood[a]	-2813.926		

However, there is no formal test in SPSS for over-dispersion (SPSS Incorporated, 2008). Dallal (2008) said “the effect of over-dispersion is to say, your point estimates are accurate but they are not as precise as you think they are”. That means the Poisson regression coefficients are correct, but the confidence interval (CI) may be weaker than stated.

In addition, the quantile classification method was used to divide the data into five groups. Table 4.7 shows the results of mean and variance for each group. Only group 5 may have problems with over-dispersion due to some extreme outliers. Therefore, 80% of data doesn’t have an over-dispersion problem. The coefficient for the group 5 (20% of data) is still considered accurate, only the confidence interval may be wider than 95%. Although, 20% of the data (group 5) have this limitation, this method is still valid for analysis of the remaining 80% of the data.

Table 4.7 Quantile classification

# of groups	# of Participation Households	Mean	Variance
1	0-3	1.548	1.343
2	4-7	5.528	1.255
3	8-11	9.429	1.239
4	12-17	14.152	2.655
5	18-66	25.482	78.068

4.5 Summary

This chapter has presented a statistic analysis of the relationship between each of the explanatory variables and the REEP participation in Waterloo Region. The explanatory variables identified from the literature review that may influence participation for this program has been used. At the end, the data limitations section was presented.

The analysis in this chapter focused on interpretation of Poisson regression model used to explore the research hypothesis that relationships exist between the REEP participation and the explanatory variables. Although the over-dispersion problem has been discussed in the limitation section, this problem may only effect on CI not coefficient, the results show that the Poisson regression model can still be considered as an appropriate method for this analysis. Based on the results obtained from Poisson regression and the correlation test, the results show all the explanatory variables were strong in predicting the relationship between each explanatory variable and participation in the REEP and they all have positive relationships as well. At the same time, according to reviewed literature the variables such as level of education, home ownership, total household income, resident age and age of dwelling all have significance in determining household energy conservation behavior.

Therefore, the conclusion can be made that analysis of all the variables became important for predicting the relationship between each variable and participation in the REEP. The next chapter will discuss another regression model to analyze the factors that may influence whether or not households will participate in the second evaluation of the REEP.

Chapter 5 Logistic Regression Analysis

5.1 Introduction

This chapter employs the third method in this research, logistic regression, to analyze the relationship between participation in the follow-up REEP evaluation and hypothesized explanatory variables. Firstly, this chapter discusses why this method was selected for the analysis; the second section presents the results of the analysis; section three discusses the limitations; and section four summarizes the conclusions.

5.2 Data Description

For binary logistic regression analysis, participating households were divided into two groups: group (0) represents households who only did the initial evaluation, and group (1) represents households who also complete the second evaluation.

In the REEP database, there was no information about the education, income, or age of people in the participating households. Therefore, CDA level data were used to represent these attributes. The use of CDA level data in this analysis is not ideal and risks producing invalid results because of the ecological fallacy problem. The ecological fallacy is a widely recognized error in the interpretation of statistical data in an ecological study, whereby inferences about the nature of individuals are based solely upon aggregate statistics collected for the group to which those individuals belong (Liggett & Hollis,1982). Liggett & Hollis (1982) show that the use of aggregated data to estimate individual level relationships is quite common in the social and behavioral science. However, there is a potential problem whenever aggregate level data are used to estimate the strength of an individual (disaggregate) level relationship. Sometimes, the aggregate level data will systematically misrepresent the true strength of the individual level relationship. In effect, the use of

CDA level data means that the statistical models are predicting household behavior based on characteristics of the neighborhood in which the household lives, rather than based on the characteristics of the household itself. This may weaken expected relationships if the characteristics of the household differ from those of the surrounding neighborhood.

The CDA attributes were linked to the REEP follow-up data by performing a spatial join. In the output shapefile there were 7214 households; 96 of which could not be matched during the geocoding process and so were not included in the analysis. For analysis of the REEP follow-up data, households who may not have had time to do a follow-up evaluation were removed from the sample. Usually, households have 18 months to do the second evaluation. Thus households who completed the initial evaluation in 2005 or 2006 may intend to do a follow-up evaluation but haven't yet had time to do it. This makes it impossible to accurately classify these households into the follow-up or non-follow up evaluation groups. To avoid possible edge effects, only households who completed initial evaluations between 1999 and 2004 were included in this analysis. These households were selected by doing a select by attributes and exporting the selected households to a new shapefile which was used as input for the statistical analysis

The data set for this analysis has a binary response (dependent or outcome) variable called **evaluation**; the total number of participating households was 5717 for this part of the analysis. There are six predictor variables: population with higher education (**bache_or_h**), average household income (**avhhinv**), population aged 65 years and over (**over65**), house size (**area_m2**), age of home(**age_home**), and increased rating score (**U_A_rating**). The first three variables (**bache_or_h**, **avhhinc**, and **perover65**) were based on information abstracted from the CDA level, and the other three (**area_m2**, **age_home**, and **U_A_rating**) were based on each REEP household (see Table 5.1).

In chapter 4, the explanatory variables were expressed as percentages while the explanatory variables in this chapter appear as counts, because this chapter focuses on individual household analysis, the use of counts was better suited than percentages in terms of interpretation of the results. In this chapter, the explanatory variables selected were based on the literature review and data availability. The literature review has discussed the variables such as household income, education level, age of dwelling, age of residence and house size all have significance effects on household energy conservation behavior.

Table 5.1 Name and Description of Explanatory Variables

1.	Average households income (CDA level)
2.	Number of Population with Higher Education (CDA level)[a]
3.	Number of Population Aged 65 Years and Over (CDA level)
4.	Age of home(Individual level)
5.	Increased rating score (Individual)
6.	Area (m^2)-home size (Individual level)

[a] attaining at least a university degree

5.3 Methodology

Logistic regression is a method for modeling the dependence of a binary response variable(also called dependent variable or outcome) on one or more explanatory variables (or predictor variables) (Bewick, Cheek & Ball, 2005; Friendly & Wuensch,1995). Generally, the dependent variable in logistic regression is dichotomous, such as response/no response or success/failure, that is, the dependent variable can use 1 to represent the response or success, and use 0 to represent no response or failure (Dobson, 1991). This type of variable is called a binary variable. Logistic regression makes no assumptions about the distribution of the predictor variables, therefore when the predictor variables are not nicely distributed the logistic regression can still be employed.

In this research, the interest is in variables that influence whether or not a household did a second REEP evaluation after upgrading their house. The dependent variable was a binary variable (0/1): zero represents the households who only did the first evaluation, and one indicates households who did the first evaluation and the second follow-up evaluation as well. The predictor variables are listed in Table 5.1. Binary logistic regression was adopted for this analysis because can handle 0/1 variable correctly and test for relationships between a binary response variable and a set of predictor variables.

Usually a regression model requires a continuous dependent variable (denoted Y). However, in this case the Y is discrete with only two categories: REEP participation households either did the second (follow-up) evaluation or not. The logistic regression turns the discrete dependent variable into a continuous output by calculating the probability *p* for the occurrence of a specific event (Menard, 2002). That means, the logistic regression provides a model to predict the *p* for a specific event for Y (i.e., p = P[Y=1]) given any value of X (here, the explanatory variables in Table 5.1).

In logistic regression, the dependent variable is a logit (a logit is a log of odds and odds are a function of P), which is the natural log of the odds, that is,

$$\text{Log (odds)} = \text{logit (P)} = \ln(P/(1\text{-}P)) \quad \text{(eq 5.1)}$$

The logistic regression equation has the form (Menard, 2002):

$$\ln (P/1\text{-}P) = \beta_0+\beta_1X_1+\beta_2 X_2\ldots+ \beta_n X_n \quad \text{(eq 5.2)}$$

Note: "The natural logarithm of the odds, ln {P(Y=1)/ [1-P(Y=1)]}, is called the logit of Y" (Menard, pp13, 2002)

People like to talk about probabilities more than odds. To get from logits to probabilities, we first have to take the log out of both sides of the equation. Then we have to convert odds to a simple probability:

The odds prediction equation: ODDS= e^{a+bx} (eq 5.3)

The convert odds to probabilities: Y= ODDS/ (1+ODDS) = e^{a+bx} / (1+ e^{a+bx}) (eq 5.4)

In SPSS, the logistic regression procedure models 1 as the response, treating 0 as the reference category. This allowed identifying which independent variables are more likely to influence households' decisions to participate in the second evaluation. The link function is logit for logistic regression. Pampel (2000, pp. 35-38) provide an example with commented SPSS output.

The **B** column in Table 5.2 displays the estimated logistic regression coefficients for the model. These are the values for logistic regression equation for predicting the dependent variable from the independent variables. They are in log-odds units. The prediction equation is

$$\log (p/1-p) = b0+b1^* x1+ b2^* x2+ b3^* x3+ b4^* x4+ b5^* x5+ b6^* x6 \quad \text{(eq 5.5)}$$

where p is the probability of doing a follow-up REEP home evaluation after implementing recommended upgrades. Expressed in terms of the variables used in this analysis, the logistic regression equation is

Log(p/1-p)= -10.65+0.000005* avhhinc-0.003* bache_or_h+ 0.002* over65+ 0.096* u_a_rating-0.003* area_m2+ 0.005* built

Recall that the dependent variable is Evaluation = 1 (the participating households did the second evaluation). These estimates tell the relationship between the independent variables and the dependent variable, where the dependent variable is on the logit scale. These estimates tell the amount of increase (or decrease, if the sign of the coefficient is negative) in the predicted log odds of Evaluation = 1 that would be predicted by a 1 unit increase (or decrease) in the predictor, holding all other predictors constant.

Because these coefficients are in log-odds units, they are often difficult to interpret, so are usually converted into odds ratios. The SPSS output lists each coefficient as an exponent in the "**Exp (B)**" column. These are the odds ratios for the predictors. They are the exponentiation of the coefficients. Pampel (2000, pp. 35-38) "states that subtracting one from each exponentiated coefficient and multiplying by 100 shows the percentage change in the odds of Evluation=1 for a one unit change in X". The **Wald** and **Sig.** column in Table 5.2 provide the Wald chi-square value and 2-tailed p-value used in testing the null hypothesis that the coefficient (parameter) is 0.

Table 5.2 Analysis of Parameter Estimates

Variables in the Equation							
	Parameter	**B**	S.E.	Wald	df	**Sig.(p-value)**	**Exp(B)**
Step 1	AVHHINC	**.000**	.000	21.076	1	**.000**	**1.000**
	BACHE_OR_H	**-.003**	.001	20.123	1	**.000**	**.997**
	OVER65	**.002**	.001	17.947	1	**.000**	**1.002**
	U_A_RATING	**.096**	.005	309.868	1	**.000**	**1.101**
	AREA_M2	**-.003**	.001	20.574	1	**.000**	**.997**
	AGE_HOME	**.005**	.002	8.969	1	**.003**	**1.005**
	Constant	**-11.732**	3.117	14.164	1	**.000**	**.000**

5.4 Results and Discussion

5.4.1 Results

Based on Table 5.2, it is interesting to note that all variables in this chapter of the study were significant in determining the number of eligible households that participate in the second evaluation. For example, the increased rating score (**U_A_RATING**): this is the logistic regression estimate for a one unit increase in rating score, given the other variables are held constant in the model. The coefficients show that one point increase in rating score increase the logged odds of **Evaluation =1**(participant households that completed the second evaluation) by 0.096, holding all other predictors constant. For the variable **U_A_rating**, with a *p*-value less than 0.0001, the null hypothesis that the coefficient equals 0 would be rejected, and we conclude that the logistic regression coefficient for **U_A_rating**, is statistically different from zero given other variables in the model.

Since the coefficients for average household income **(AHHINC**), aged population (**OVER65**)**,** and age of home (**age_home**) are positive, similar interpretations can be made for these variables, although the increases in logged odds of Evaluation=1 would be smaller because these coefficients are smaller.

The coefficients for dwelling size (**AREA_M2**) and education (**BACHE_OR_H**) are negative and both are equal to -0.003. Therefore, the same interpretations can be made for these two variables as well. For example, the dwelling size: this is the logistic regression estimate for one unit increase in dwelling size, given the other variables are held constant in the model. The coefficients show that one square meter increase in dwelling size lower the logged odds of **Evaluation =1** by0.003, holding all other predictors constant. For the variable dwelling size, with a *p*-value less than 0.0001, the null hypothesis that the coefficient equals 0 would be rejected, and we conclude that the logistic regression coefficient for dwelling size is statistically different from zero given other variables in the model.

The p-values indicate that all predictor variables are significant at the 95% confidence level. Thus, in each case, we would reject the null hypothesis of no relationship between the independent and dependent variables when the other variables in the model are held constant.

5.4.2 Discussion

The dependent variable indicates whether or not households did the second evaluation, and all the explanatory variables have a highly significant effect on the outcome (*p* less than 0.0001, except age of house with p equal to 0.003).

The resulting average household income coefficient of 0.005 shows that one thousand dollars increase in average household income increase the odds of household participation in second evaluation by 0.05%. The result shows that higher income is associated with higher participation in the second evaluation. The second evaluation (also called follow up evaluation) requires certain costs. Households who participated in the second evaluation may already have invested money to retrofit their homes. Households with higher income may better afford this investment and therefore be more likely to participate in the follow-up evaluation.

Surprisingly, population with higher education was found to have a negative effect on household participation for the second evaluation. The resulting coefficient of -0.003 shows that estimated odds of household participation in the second evaluation decrease by 0.3% for every increase in number of people with higher education by one person. That may be due to higher educated people having a good job and competitive salary, and also may be too busy to participate. They don't care how much money can be saved from this program and they may not have time to do this evaluation as well. However, the first evaluation found that higher educated people make a great contribution to participation in the energy efficiency project. In the literature, a lot of researchers found that people with higher education had a better understanding of energy conservation, and also the people with more education may have greater concern about the seriousness of the energy situation. But other studies found that there is no relationship between energy use and education level.

The possible explanation for the relationship between aged population and household participants for the second evaluation may be because some older people will stay home longer than younger adults. During the winter, the older adults used more electricity than younger adults on average. This may indicate that older adults have a lower tolerance for cold, which causes additional heating requirements. The situation mentioned above has been proved in the literature review. These situations can indicate older adults may use more energy than younger adults, therefore, if they did retrofit their homes, they can have a great energy savings from it. Moreover, the estimated odds of household participation in the second evaluation increased by 0.2% for every increase in number of population aged 65 years and over by one person.

The rating score was also significantly associated with household participation in the second evaluation. The resulting coefficient of 0.096 shows that a one point increase in rating score increases the odds of household participation in the second evaluation by 10.1%. The rating score measured the efficiency of homes, ranging from 0 to 100. If there is a greater difference between evaluation one and two, this indicates that there was more opportunity to upgrade the homes and upgrading may increase the resale value. Usually, energy efficient homes have resale values that are 2%-8% higher than less efficient homes (REEP, 2007a). People living in relatively low energy efficiency homes may be more likely to participate in this evaluation because they can easily save money by doing some improvements for their homes, and also gain government grant.

The result of this analysis shows that the dwelling size has a negative effect on outcome. The coefficient of -0.003 shows that estimated odds of household participation in the second evaluation decreased by 0.3% for every one square meter increase in dwelling size. A logical explanation of this relationship may be due to the retrofit costs for large homes are more expensive than for small homes, so owners of larger homes may decide not to participate.

The current study also found that the age of home is significant in determining participation in the second evaluation. However, this result shows that an older home may have less incentive to participate in the second evaluation. The estimated odds ratio is very close to 1. It turns out that estimated odds of household participation in the second evaluation increase by 0.5% for every one year increase in age of the house. Older homes may require larger investment for retrofit, so the owner may be unwilling to make a large investment unless they plan to remain in the same house for several years. They may decide to perform only a portion an upgrade for their homes that the report suggested, but didn’t come back for the second visit.

5.5 Model Validation

In order to validate the logistic model, the receiver operating characteristic (ROC) curve was investigated. This procedure is a useful way to evaluate the performance of classification schemes in which there is a variable with two categories by which subjects are classified (SPSS Inc., 2008).

Table 5.3 represents the results obtained from ROC curve. The area under the ROC curve is a number between 0.5 to 1. The area equal to 0.5 indicates no separation of the two classes, and the area equal to 1 indicates perfect separation of the two classes. The results show in Table 5.3 the area under the curve is 0.746 and that shows the logistic model did well predicting which households completed the second evaluation. Therefore, the results from the logistic model are reliable.

Table 5.3 Areas under the ROC Curve

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.746	.007	.000	.732	.761

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5
Test Result Variable(s):Predicted probability

5.6 Summary

This chapter has presented a logistic regression model for the analysis of the relationship between each explanatory variable and whether or not to participate in the second evaluation. The selected explanatory variables were identified from the literature review and also based on data availability.

The results show all the explanatory variables have significant relationships with the dependent variable. The increased rating score (**U_A_RATING**), average household income **(AHHINC**), aged population (**OVER65),** and age of home (**AGE_HOME**) are positively related to the dependent variable. While the dwelling size (**AREA_M2**) and education (**BACHE_OR_H**) has negative relationships with the dependent variable.

Based on reviewed literature the variables such as level of education, dwelling size, total household income, resident age and age of home all have significance in determining household energy consumption behavior. However, the results of the statistical analysis indicate that although the model does a good job of predicting which households will not participate in the follow-up evaluation, it did a poor job of predicting which households would participate. The predictive ability of the model may have been weakened because of the use of CDA level data to represent household characteristics such as age, education and income that were not available in the REEP dataset. There may also be other factors that affect behavior that were not included in the analysis.

Chapter 6 Conclusions and Recommendations

6.1 Conclusions

The main objectives of this thesis are:

1. To understand household behavior toward energy conservation;
2. To explore the spatial distribution of participating households each year and overall and assess the marketing methods that play an important role in influencing participation;
3. To determine the factors affecting participation in the residential energy efficiency program in the Waterloo Region;
4. To make recommendations to encourage greater participation in future programs of this type, and for future research.

The main conclusions of this study are summarized in the following paragraphs.

From the literature review, variables such as resident age, home ownership, education level, total household income, number of occupants, house size, and age of the home all affect energy consumption to a varying degree. The literature review aided in determining what variables should be included in this research.

GIS analysis was used to present a visual interpretation of spatial patterns of REEP participation in Waterloo Region, the impact of sources of information on participation and the relationship between participation rates and explanatory variables.

The spatial pattern of participation changed over time (from 1999 to 2006). In general, participating households were clustered in urban areas, and other small communities like Elmira, Wellesley, Baden, New Hamburg, New Dundee, and Ayr; and were scattered in rural areas.

Participating households first developed from the City of Waterloo and City of Kitchener, and then expanded to the City of Cambridge, to form a very dense pattern through eight years. Furthermore participation seems to expand to southeast first, and then southwest, followed by the northwest and northeast areas.

The different sources of information came to play a significant role in different years. In 1999, most evaluations originated from media (45.7%), followed by referral (31.3%), community-based event (8.7%), community based meeting (6%), direct marketing (4.9%) and passive marketing (0.38%). REEP relied heavily on media and referrals during this year. However, after 1999 the direct marketing increased from 2000 to 2002, and decreased from 2003 due to other sources of information, such as community-based event and passive marketing becoming more important. Referrals have remained an important source of information for evaluations and increased in percentage from 2003 to 2005.

Finally, Spearman's rank correlation was used to assess the strength of associations between hypothesized explanatory factors (level of education, average household income, employment rate, home ownership, population aged 65 and over, age of dwelling, and number of eligible dwellings) and participation rates. In general, these results confirmed the findings based on map interpretation and indicate moderate positive associations between participation rates and each of the explanatory factors.

Statistical analysis, using a Poisson regression model to determine the relationship between the independent variables and the dependent variable, showed approximately one-third of the variation in participation could be explained with all independent variables. These variables include level of education, average household income, employment rate, home ownership, population aged 65 and over, age of dwelling, and number of eligible dwellings. The most important determinant of households choosing to participate in the REEP was people with higher education; followed by participants who own their home and age of dwellings. A one percent increase in the percentage of people having higher education was expected to increase the log of the number of households participating in REEP by 0.29. Similarly, a one percent increases in the percentage of people who own homes and dwellings built before 1970 was expected to increase the log of the number of households participating in REEP by 0.08 and 0.06 respectively. The direction of each of these relationships supports the findings in previous literature.

Diagnostic tests revealed that there may be an over-dispersion problem for the Poisson regression model. However, further investigation showed that this may only be the case for the top 20% of data and may be caused by some extreme outliers. This problem primarily affects the accuracy of confidence intervals not the coefficients. It was concluded that the Poisson Regression model is appropriate for the data used in this analysis and that the results represent real relationships between the hypothesized explanatory variables and participation in the REEP.

Results of the logistic regression analysis were used to determine the relationship between the explanatory variables and a binary dependent variable that indicated whether or not households completed the second evaluation. All the explanatory variables (household income, level of education, population aged 65 and over, age of home, rating score, and house size) had a significant effect on the outcome (*p* less than 0.0001, except age of home with p equal to 0.003). Although this model was able to correctly predict the behavior of 97.7% of households who did not participate in the second evaluation, it was only able to correctly predict 7.7% of households who did participate. The weak predictive ability for the latter group may be due in part to the use of aggregate CDA level data to represent household characteristics that were not available in the REEP data. There may be other factors not included in this analysis that might improve the predictive ability of the model, but these results may simply reflect the randomness of household decision making.

The increased rating score, average household income, population aged 65 and over and age of dwelling all have positive relationships with the dependent variable. The most important determinant of households choosing to participate in the second evaluation was an increased rating score. A one point increase in rating score increases the odds of household participation in second evaluation by 10.1%. While, the coefficients for house size and education are negative and the same, this means they have a negative relationship with the dependent variable. The coefficient of -0.003 shows that estimated odds of household participation in the second evaluation decreased by 0.3% for every square meter increase in dwelling size, and for every increase in the number of people with higher education as well.

6.2 Research Contributions

Understanding residential energy consumption behavior is necessary for successfully promoting residential energy efficiency programs. A modest attempt was made to improve this understanding with this research. General speaking, this research has focused on identifying the variables affecting the REEP participation, and exploring the changes of the spatial pattern through 1999 to 2006and explaining how this pattern happened, and why there is a relationship between dependent variable and independent variables. It was hoped that this study would contribute to the growing body of knowledge on the residential energy efficiency program in the following ways.

In general, this work has provided a practical understanding of how the energy efficiency program operates, and provides insight into the type of variables that may be successful in bringing about changes in performance in the energy efficiency project in Waterloo Region. It also provides insight as to whether the selected variables affect consumers' attitudes towards energy conservation issues, which may help to further motivate households to exhibit conservation behavior. In terms of contribution to the literature review, this research confirmed what has been leaned from the literature review. Variables affecting the energy consumption behavior in past still affect in present (refer section 2.2, 4.3 and 5.4). The barriers to participation in the energy efficiency program such as the indifferent, cost sensitive, or lacking in full knowledge in residential energy efficiency program are still the barriers for present to attract more participation. Except cost sensitive, this research made recommendation for other barriers for future success in this kind of programs. The cost sensitive only can solve through monetary support from varies organizations or government, which cannot force them to do that. It's also dependent on the priority of cities or regions development direction. If a city focuses on environmental oriented development, it may invest a great amount of money to encourage the programs are environmental friendly.

Secondly, with the completion of this research, future residential energy efficiency program can use the information from this research and emulate or expand upon the efforts and lessons learned from the Residential Energy Efficiency Project in Waterloo Region case example. Other residential energy efficiency programs can learn how these identified explanatory variables affect the participation and these variables (also reflected the consumers concerns) can be integrated into their residential energy efficiency program design plans and improve the level of participation rate in their area when the program starts because sometimes the participation rate in the first year may lower than following years.

Thirdly, this research also contributes to a practical experience on how to integrate different datasets using GIS. GIS provides a platform to visualize valuable physical and socio-demographic data of the neighborhood in the program, including GIS as well as various integrated data stores. Furthermore, this application also can be used in the other energy data and business data. This allows utilities to leverage the power of GIS to add meaning that transforms distributed data into information. In terms of energy efficiency analysis GIS still in a new role, not a lot of researchers use this technique to evaluate or analyze how well the program was developed. This research may contribute to other organizations that are interested in learning how to use GIS to do the overall program evaluation and factors analysis.

In terms of statistical analysis, the Poisson and binary logistic regression models used in this research have not been widely used in describing or predicting participation in energy efficiency programs. From an academic perspective, this research contributed empirical evidence to the bodies of work describing various methods such as two statistical methods. From practical perspective, this research provided an alternative way to analyze factors expected to influence participation in energy efficiency programs. That may help other researchers who have similar to find a better way to analyze the causes and to predict effects. At the same time, this research provides examples for other readers who may be interested in using Poisson or logistic regression in SPSS, and may help them learn how to interpret the results.

In terms of contribution to the human ecosystem model, this research found the human environment had the most affect on participation in energy efficiency, the social environment such as social norms and designed environment have some affect on participation in energy efficiency as well, but not as much as the human environment. Since the research only included two variables in the designed environment, other variables should be examined in future research. Also this research did not include any social environment variables in the statistical test due to the lack of this type of data.

6.3 Recommendations

Continued research on residential energy consumption should occur. The current descriptive findings provide insight into variables at the aggregate level and household level that affects participation in the REEP. From the literature review and results of the present study, the following areas of future research are suggested:

Research on other explanatory variables that may affect energy consumption behavior is necessary to be able to better predict future participation in this kind of program, and especially to predict participation in the second follow-up evaluation. If the other explanatory variables can be found that influence the REEP participation, that may help to improve the overall quality of the Poisson regression and logistic regression models (refer to section 4.4 and 5.5).

Conducting interviews based on households who already participate in the REEP may aid in the collection of individual level data as opposed to using aggregate level of data, thus avoiding the problem of ecological fallacy.

To improve upon the energy efficiency program, an energy education program should be created. That can help improve people's awareness of the energy efficiency program and increase the probability of participation in the energy efficiency program. As the energy efficiency program continues, three parts including people, message, and longevity of the program should be considered to ensure the program success.

People need to be aware of the energy efficiency programs, whether they are new or existing households in their area, and have an accurate understanding of the potential savings from that program. This is most effectively achieved by targeting the individual household, based on characteristics identified in this study, rather than sending out general tips in a brochure. For example, based on the results obtained from this research, higher income and higher education people are more likely to participate in the energy efficiency program. If the program can use some specific marketing method to develop these groups of population in each area then they may get more participating households.

The program must identify what message will motivate occupants to change behavior patterns effectively and provide useful, frequent feedback to encourage participation for households in the energy efficiency program. If the local utility company can get involved, and they provide detailed information for each household's energy consumption based on monthly billing, and also include the previously energy consumption from last month, that will give each household frequent feedback on their energy bill and may attract them to pay attention to use energy in a more efficient way. Therefore, when the local residences know they have an energy efficiency program, they may show more interest to participate in the program.

Finally, the energy efficiency program needs longevity, evidenced by sticking with its goals and encouraging continued participation. Usually, long lasting program will become popular in local areas and gain more participants. Consequently, long-running program is very important for future energy efficiency program success. This may also be why the REEP still used the same name as before, because well known of the project.

The last recommendation addresses people who are indifferent about residential energy efficiency. These consumers may be aware of energy options, but still make inefficient energy-related choices. Therefore, if this problem can be solved then this part of group people will participate. As Reddy (1991) suggests government intervention through regulations, standards, restriction in supply are still good to do that for present.

References

Anderson, C. D., & Claxton, D. J. (1982). Barriers to consumer choice of energy efficient products. *The Journal of Consumer Research, 9* (2), 163-170.

Berk, R. & MacDonald, J. (2007). *Overdipersion and poisson regression*. Retrieved August 20, 2008, http://www.crim.upenn.edu/faculty/papers/berk/regression.pdf

Brandon, G., & Lewis, A. (1999). Reducing household energy consumption: a qualitative and quantitative field study. *Journal of Environmental Psyhology, 19* (1), 75-85

Berry, L., Soderstrom, J., Hirst, E., Newman, B., & Weaver, R. (1981). *Review of evaluations of utility home energy audit programs* (NTIS No: ORNL/CON-58). Oak Ridge, TN:Oak Ridge National Laboratory.

Bubolz, M., Eicher, J., & Sontag, M. (1979). The human ecosystem: a model. *Journal of Home Economics*, 28-31.

Black, S. J., Stern, P. C., & Elworth, J. T. (1985). Personal and contextual influences on household energy adaptations. *Journal of Applied Psychology*, *70* (1), 3-21.

Bewick, V., Cheek, L., & Ball, J. (2005). Statistics review 14: logistic regression. *Critical Care*, *9,* 112-118.

Brown, M., Berry, L., Bdzer, R., &Faby, E. (1993). *National impacts of the weatheriation assistance program in single-family and small multifamily dwellings.* Oak Ridge National Laboratory, ORNL/CON-326.

Brown, MA. (2001). Market failures and barriers as a basis for clean energy policies. *Energy Policy, 29*, 1197-1207.

Cohen, S., Goldman, C., &Harris, J. (1991). *Measured energy savings and economics of retrofitting existing single-family homes: an update of the BECA-B database.* Lawrence Berkeley Laboratory LBL-28147 UC-350 (volume two).

City of Toronto. (2008). *Energy efficiency office.* Retrieved October 30, 2008, from http://www.toronto.ca/energy/

Clark, K.E., & Berry, D. (1995). House characteristics and the effectiveness of energy conservation measures. *Journal of the American Planning Association, 61*(3), 286-359.

City Green Solutions. (2008). *One application process. Two grant cheques!* Retrieved May 20, 2008, from http://www.citygreen.ca/energy/ecoENERGY.aspx

Cameron, A.C., and Trivedi, P.K. (1998). *Regression analysis of count data.* Cambridge: Cambridge University Press.

City of Vancouver. (2008). *Incentive programs.* Retrieved October 20, 2008, from http://vancouver.ca/oneday/takeAction/atHome/incentive.htm

Dallal. G.E. (2008). *Poisson regression.* Retrieved August 20, 2008, from http://www.jerrydallal.com/LHSP/Poisson.htm

Dobson, A.J. (1991). *An introduction to generalized linear models.* New York: Wiley.

EnviroCenter. (2007). *Free energy conservation program in Ottawa.* Retrieved November 1, 2007, from http://www.envirocentre.ca/envirocentre_eng.html

Energy Savings Plan. (2007). *The benefits of energy efficiency.* Retrieved November 1, 2007, from http://www.saveenergynow.ca/caseEE

Emery, A. F., & Gartland, L. M. (1996). Quantifying occupant energy behavior using pattern analysis techniques. In *Proceedings of the ACEEE, 1996 summer study on energy efficiency in buildings 8*, 47-59.Washington, DC: American Council for an Energy Efficient Economy.

Farthar, BC. (1994). Trends in U.S. public perceptions and preferences on energy and environmental policy. *Annual Review of Energy and the Environment, 19*, 211-239.

Farthar, B.C. & Houston, A. H. (1999). *Willingness to pay for electricity from renewable energy*. Golden, Colorado: A national laboratory of the U.S. department of energy managed by midwest research institute for the U.S. department of energy.

Friendly, M. (1995). *Logistic regression*. Retrieved August 2, 2008, from http://www.math.yorku.ca/SCS/Courses/grcat/grc6.htm

Gatersleben, B., Steg, L., & Vlek, C. (2002). Measurement and determinants of environmentally significant consumer behavior. *Environment and Behaviour, 24* (3), 335-362.

Guerin, D.A., Yust, B. L., & Coopet, J.G. (2000). Occupant predictors of household energy behavior and consumption change as found in energy studies since 1975. *Family and Consumer Science Research Journal, 29* (1), 48-80.

Gmelch, L., & Dillman, J. (1988). House hold energy conservation motivators: A factor analytic approach. *Housing and Society*, *13*(3), 238-245.

Green Communities Canada. (2008). *Backgrounder-energy efficiency assistance program for houses.* Retrieved August 2, 2008, from http://www.gca.ca/indexcms/downloads/EEAPH_Backgrounder.pdf

Guerin, D. (1992). Interior design research: a human ecosystem model. *Family and Consumer Science Research Journal*, *20*(4), 254-263.

Hagelstam, H., & Hollins, J. (2002). *Sustainable energy: engaging citizens*. Energy Council of Canada. Retrieved August 2, 2007, from http://libdigi.unicamp.br/document/?view=17

Harchaoui, T. (2003). *Greenhouse gas emissions in the Canadian economy, 1981-2001*. Analytical Paper: Insights on the Canadian Economy. Statistics Canada Catalogue No. 11-624-MIE No. 001.

Hirst, E., & Goeltz, R. (1984). *Comparison of actual and predicted energy savings in Minnesota gas heated single family homes* (NTIS No: ORNL/CON-147). Oak Ridge, TN: Oak Ridge National Lab oratory.

International Energy Agency. (2005). *A marketing guide based on experiences from 10 countries.* Retrieved April 2, 2008, from http://www.ecbcs.org/docs/Annex_38_IEA_Sustainable_Housing_Marketing_Guide.pdf

Johnson-Carroll, K., Brandt, J., & Olson, G. (1987). Factors that influence energy conservation alterations in Oregon households. *Housing and Society*, *14*(2), 111-129.

Kitchener Utilities. (2005). *Residential energy efficiency project (REEP).* Retrieved February 20, 2007, from http://www.kitchenerutilities.ca/resource_centre/quality_conservation/reep.asp#eligible

Kempton, W., Harris, C., Keith, J., & Weihl, J. (1985). Do consummers know "what works" in energy conservation? *Marriage and Family Review*, *9*(1/2), 115-133.

Kasulis, J., Huettner, D., & Dikeman, N. (1981). The feasibility of changing electricity consumption patterns. *Journal of Consumer Research*, *8*, 279-290.

Liggett, R.S., & Hollis, M.E. (1982). *The ecological fallacy and the use of spatially aggregated data.* Los Angeles : Graduate School of Architecture and Urban Planning, University of California.

McMakin, A. H., Malone, E.L., & Lundgren, R.E. (2002). Motivating residents to conserve energy without financial incentives. Environment and Behavior, 34 (6), 848-863.

Mitchell, A. (2005). *The ESRI guide to GIS analysis: spatial measurements & statistics.* California: ESRI Press.

Meyers, S. (1998). *Improving energy efficiency: strategies for supporting sustained market evaluation in developing and transitioning countries.* Retrieved Novermber 10, 2007, from http://ies.lbl.gov/iespubs/41460.pdf

Menard, S. (2002). *Applied logistic regression analysis, 2nd Edition*. Thousand Oaks, CA: Sage Publications. Series: Quantitative Applications in the Social Sciences, No. 106. 1st ed., 1995.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models.* 2nd Edition. London: Chapman & Hal.

Natural Resources Canada (NRCan). (2007a). *The EnerGuide rating service*. Retrieved June 20, 2008, from http://www.oee.nrcan.gc.ca/residential/personal/new-homes/upgrade-packages/energuide-service.cfm?attr=4).

Natural Resources Canada. (2007b). *Directory of energy efficiency and alternative energy programs in Canada.* Retrieved January 2, 2008, from http://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/policy_e/programs.cfm?programtypes=all&categoryid=all&regionaldeliveryid=all&attr=0

National Research Council. (2005). *Decision making for the environment: Social and behavioural science and research priorities.* Washington DC: National Academy Press.

Nielsen, L. (1993). How to get the birds in the bush into your hand: Results from a Danish research project on electricity savings. *Energy Policy, 21* (11), 1133-1144.

Natural Resources Canada. (2008). *Retrofit your home and qualify for a grand!* Retrieved Auguest 2, 2008, from http://oee.nrcan.gc.ca/residential/personal/retrofit-homes/retrofit-qualify-grant.cfm?attr=4

Natural Resources Canada. (2006a). *Energy efficiency trends in Canada: 1990-2004.* Retrieved June 20, 2007, from http://www.oee.nrcan.gc.ca/Publications/statistics/trends06/pdf/trends06.pdf

Natural Resources Canada. (2006b). *The state of energy efficiency in Canada.* Retrieved June 20, 2007, from http://oee.nrcan.gc.ca/Publications/statistics/see06/pdf/see06.pdf

Natural Resources Canada. (2006c). *The news room- energy efficiency regulations.* Retrieved June 20, 2007, from http://www.nrcan-rncan.gc.ca/media/newcom/2007/200704b-eng.php

OPA (Ontario Power Authority). (2008). *OPA initiatives.* Retrieved June 20, 2008, from http://www.powerauthority.on.ca/Page.asp?PageID=122&ContentID=4172&SiteNodeID=131

Parker, P., Rowlands, I.H. & Scott, D. (2004). *Comparing residential energy conservers and consumers: local programs need all income groups to achieve Kyoto targets.* Energy Studies Working Paper.

Poortinga, W., Steg, L., & Vlek, C. (2004). Values, environmental concern, and environmental behavior- a study in to household energy use. *Environment and Behavior, 36* (1), 70-93.

Peters, J. (1990). Integrating psychological and economic perspectives of thermostat setting behavior. *In Proceedings of the American Council for an Energy Efficient Economy, 1990 summer study on energy efficiency in buildings 2,*111-119. Washington, DC: American Council for an Energy Efficient Economy.

Pampel, F. C. (2000). *Logistic regression: a primer*. Sage Quantitative Applications in the Social Sciences Series #132. Thousand Oaks, CA: Sage Publications.

Residential Energy Efficiency Project (REEP). (2007a). *ecoENERGY home evaluation. ?* Retrieved May 20, 2008, from http://www.reepwaterlooregion.ca/prog_ecoenergy.php

Residential Energy Efficiency Project (REEP). (2007b). *What is REEP?* Retrieved May 20, 2008, from http://www.reepwaterlooregion.ca/about_reep.php

Rowlands, I.H., Parker, P., & Scott, D. (2001). Citizen and *consumer attitudes towards electricity industry restructuring: an Ontario (Canada) case-study.* Energy Studies Working Paper.

Regional Municipality of Waterloo. (1998). *Region of Waterloo Statistical Profile: Agriculture.* Retrieved May 20, 2008, from http://region.waterloo.on.ca/web/region.nsf/DocID/0776E1882A72B3DC85256B1B006F8ADB/$file/StatProfile_Ag_1998.pdf?openelement

Rokeach, M. (1973). *The nature of human values*. New York: Free Press.

REEP 2006 Annual Report. (2007). Annual report. Retrieved May 20, 2007, from http://www.reepwaterlooregion.ca/about_annual_report.php

Reddy, K.N. (1991). Barriers to improvements in energy efficiency. *Energy Policy, 19*(10), 953-961.

Region of Waterloo. (2008). *About the region.* Retrieved May 20, 2008, from http://www.region.waterloo.on.ca/WEB/Region.nsf/$All/EC8014200B8AF5FA85256AF700553207?OpenDocument

Schweitzer, Martin, Eric Hirst, and Lawrence Hill. (1991). *Demand-side management and integrated resource planning: findings from survey* of *24 electric utilities.* Oak Ridge National Laboratory ORNL/CON-3 14.

Schweitzer, M., & Tonn, B. (2001). *Non-energy benefits from the weatherization assistance program: a summary of findings from the recent literature.* ORNL/CON-484. Oak Ridge National Laboratory, Oak Ridge, TN.

Stewart, K., & Fyr, K. (2006). *A low-income energy efficiency program: mapping the sector and program design principles.* A Report prepared by the Toronto Environmental Alliance for the Ontario Power Authority's Conservation Bur. Retrieved June 17, 2007, from http://www.conservationbureau.on.ca/Storage/13/1834_Low_Income_Energy_Efficiency_Program.pdf

Seligman, C., Darley, J., & Becker, L. (1978). Behavioral approaches to residential energy conservation. *Energy and Building, 1*, 325-337.

Statistics Canada. (2007). *Waterloo, Ontario* (table). *2006 community profiles*. 2006 Census. Statistics Canada Catalogue no. 92-591-XWE. Ottawa. Released March 13, 2007.

http://www12.statcan.ca/english/census06/data/profiles/community/Details/Page.cfm?Lang=E&Geo1=CD&Code1=3530&Geo2=PR&Code2=35&Data=Count&SearchText=waterloo&SearchType=Begins&SearchPR=01&B1=All&Custom

Schulze, W.D. (1994). *Green pricing: solutions for the potential free-rider problem*. Boulder, CO: University of Colorado.

Straughan, R. & Roberts, J. (1999). Environmental segmentation alternatives: A look at green consumer behavior in the new millennium. *Journal of Consumer Marketing, 16* (6), 558-575.

Strien, A.V., Pannekoek, J., Hagemeijer, W & Verstrael, T.. (2004). A Loglinear poisson regression method to analyse bird monitoring data. In: Anselin, A (ed.) Bird Numbers 1995, Proceedings of the International Conference and 13th Meeting of the European Bird Census Council, Parnu, Estonia. *Bird Census News* 13 (2000): 333-39.

Stern, C.P. (1999). Information, incentives, and proenvironmental consumer behavior. *Journal of Consumer Policy, 22*, 461-478.

SPSS Incorporated. (2008). *SPSS for Window16.0.* Chicago: SPSS Inc.

Tonn, B., & Peretz, J.H. (2007). State-level benefits of energy efficiency. *Energy Polilcy, 35*, 3665 – 3674.

Tienda, M., & Aborampah, O. (1981). Energy-related adaptations in low-income nonmetropolitan Wisconsin counties. *Journal of Consumer Research*, *8*, 265-270.

Winfield, MP., & Hall, S. (2006). *A quick-start energy-efficiency strategy for Ontario.* Retrieved November 20, 2007, from http://pubs.pembina.org/reports/quickstart_Final_Apr0606.pdf

Warriner, G. (1981). Electricity consumption by the elderly: Policy implications. *Journal of Consumer Research*, *8*, 258-264.

Wilhite, H., & Ling, R. (1995). Measured energy savings from a more informative energy bill. *Energy and Buildings, 22*,145-155.

Weihl, J. S., & Gladhart, P. M. (1990). Occupant behavior and successful energy conservation: Findings and implications of behavior monitoring. In *Proceedings of the American Council for an Energy-Efficient Economy, 1990 summer study on energy efficiency in buildings* (pp. 2.171-2.180). Washington, DC: American Council for an Energy Efficient Economy.

World Energy Council (WEC). (2008). *Energy efficiency policies around the world: review and evaluation.* Retrieved April 2, 2008, from http://www.worldenergy.org/wec-geis

Yukon Housing Corporation. (2007). 2007 Residential energy programming. Retrieved October 9, 2007, from http://www.housing.yk.ca/

Yust, B.L., Guerin, D.A., & Coopet, J.G. (2002). Residential energy consumption: 1987-1997. *Family and Consumer Science Research Journal, 30* (3), 323-349.

Zarnikau, J. (2003). Willingess to pay a premium for energy efficiency. *Energy Policy, 31*, 1661-1672.

Appendices

A. Correlation Analysis

B. Glossary

C. Summary of Sources of Information by Year

D. Q-Q Plot

E. EnerGuide: Energy Efficiency Rating

F. Summary of Statistic of Dependent Variable

G. Parameter Estimates Comparison: Weighted Dependent Variable vs. Unweighted Dependent Variable

H. Scatter Plot

Appendix A

Correlation Analysis

A correlation analysis was performed in SPSS (SPSS Inc., 2008), a statistical software package. The count value for each variable and number of eligible households in each CDA were used as input into the analysis. The results of the correlation analysis are provided in the table below.

Name of Factors	Test Factors	Spearman's rho Correlation Coefficient
Population with Higher Education	**Bache_or_H**	0.454**
Average Household Income	**Avhhinc**	0.417**
eligible households (dwellings)	**Appr_Dwe**	0.405**
Employment Rate	**Employ_Rte**	0.093*
Dwellings Occupied by Owner	**Owned**	0.453**
Population Aged 65 Years and Over	**Over65**	0.238**
Dwellings Built Before 1970	**Bf1970**	0.186**

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

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

Key to interpreting the Spearman's rho Correlation Coefficient:

1) Strength of the relationship is determined by the value of the coefficient (r)

$-0.3 \leq r \leq 0.3$ indicates a weak relationship

$-0.7 \leq r < -0.3$ or $0.3 \leq r < 0.7$ indicates a moderate relationship

$-1.0 \leq r < -0.7$ or $0.7 < r \leq 1.0$ indicates a strong relationship

2) Direction of the relationship is determined by the sign of the coefficient (r)

$r = 0$ indicates that there is no relationship

$r < 0$ indicates a negative relationship

$r > 0$ indicates a positive relationship

Appendix B

Glossary

Dissemination Area**:** Small area composed of one or more neighboring blocks, with a population of 400 to 700 persons. All of Canada is divided into dissemination areas.

Geocoding**:** The process of assigning geographic identifiers (codes) to map features and data records. The resulting geocodes permit data to be linked geographically.

Thematic Map: A thematic map shows the spatial distribution of one or more specific data themes for standard geographic areas. The map may be qualitative in nature (e.g. predominant farm types) or quantitative (e.g. percentage of population changes).

Classification Method: There was several classification methods can be used when classes group features with similar values, by assigning same symbol with similar values. There are five method to create classes include creating classes manually, using Natual Breaks (Jenk's), Quantile, Equal Interval, or Standard Deviation. In this research, when each thematic maps were create for each test factors, the Quantile classification method was used for all test factor's thematic maps, in order to do spatial analysis. Since the same classification method was easy for compare the difference between each test factors and the Quantile classification method has an equal number of features in each class that allowed later on to calculate the percentage of overlapped area (Mitchell, 1999).

Spatial Data: Information (data) that contains spatial elements. The minimum spatial information is generally considered to be location, which is typically expressed using either Eastings and Northings or Latitude and Longitude, although there are many other coordinate systems. Spatial Data may also contain information about the shape and size (geometry) and relationships to other entities.

Generalized Linear Model (GML): In statistics, the generalized linear model is a flexible generalization of ordinary least squares regression. It relates the random distribution of the measured variable of the experiment (distribution function) to the systematic (non-random) portion of the experiment (the linear predictor) through a function called the link function. Generalized linear models were formulated by John Nelder and Robert Wedderburn as a way of unifying various other statistical models, including linear regression, logistic regression, and Poisson regression, under one framework. This allowed them to develop a general algorithm for maximum likelihood estimation in all these models. It extends naturally to encompass many other models as well.

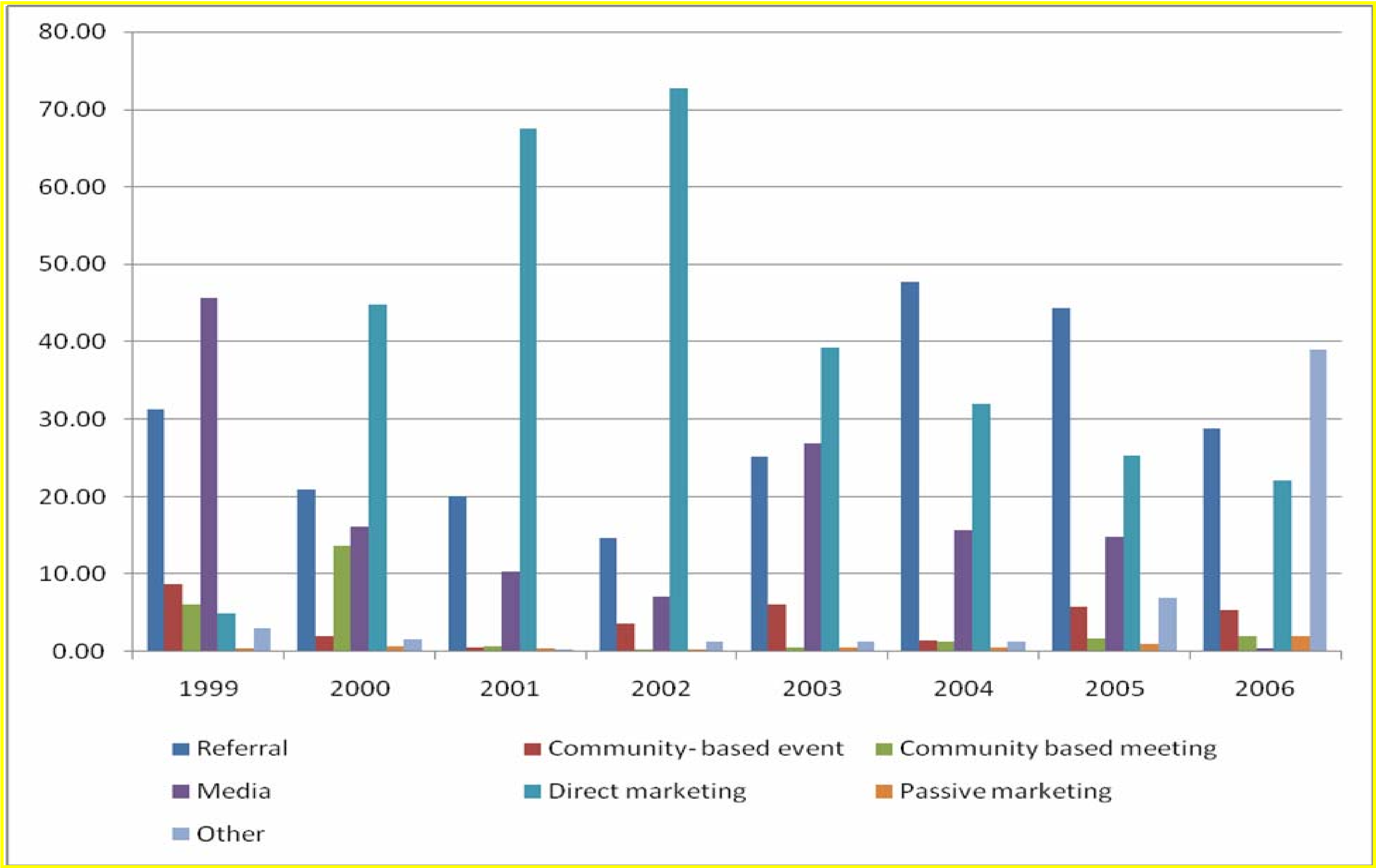
Poisson distribution: The Poisson distribution arises when you count a number of events across time or over an area. You should think about the Poisson distribution for any situation that involves counting events.

Energy Conservation: Energy Conservation is the practice of decreasing the quantity of energy used. It may be achieved through energy efficiency use, in which case energy use is decreased while achieving a similar outcome, or by reduced consumption of energy services.

Energy Efficiency: Efficient energy use, sometimes simply called energy efficiency is using less energy to provide the same level of energy service.

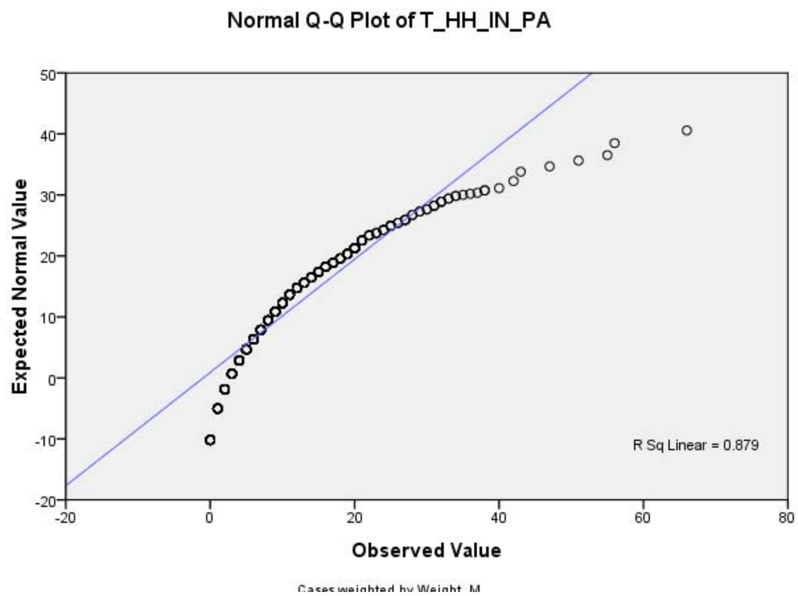
Appendix C

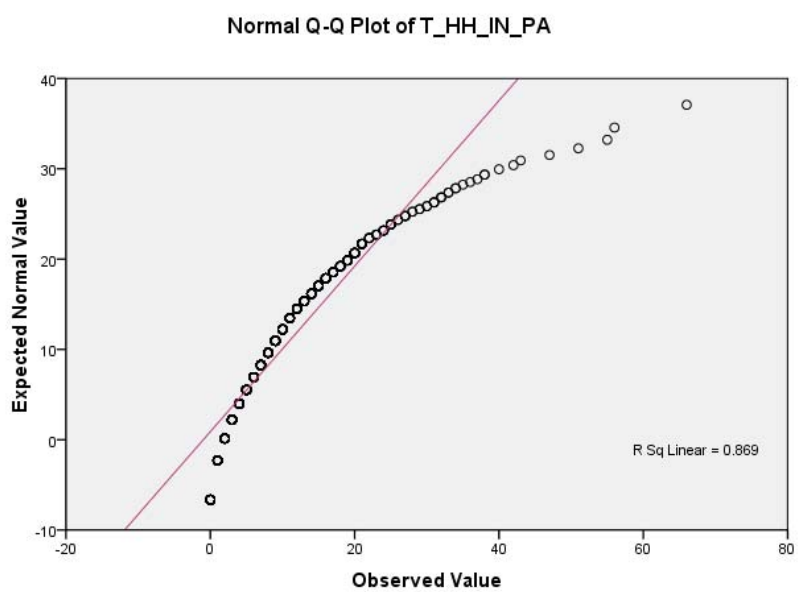
Summary of Sources of Information by Year


80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
1999
2000
2001
2002
2003
2004
2005
2006
Referral
Community- based event
Community based meeting
Media
Direct marketing
Passive marketing
Other

Appendix D

Q-Q plot

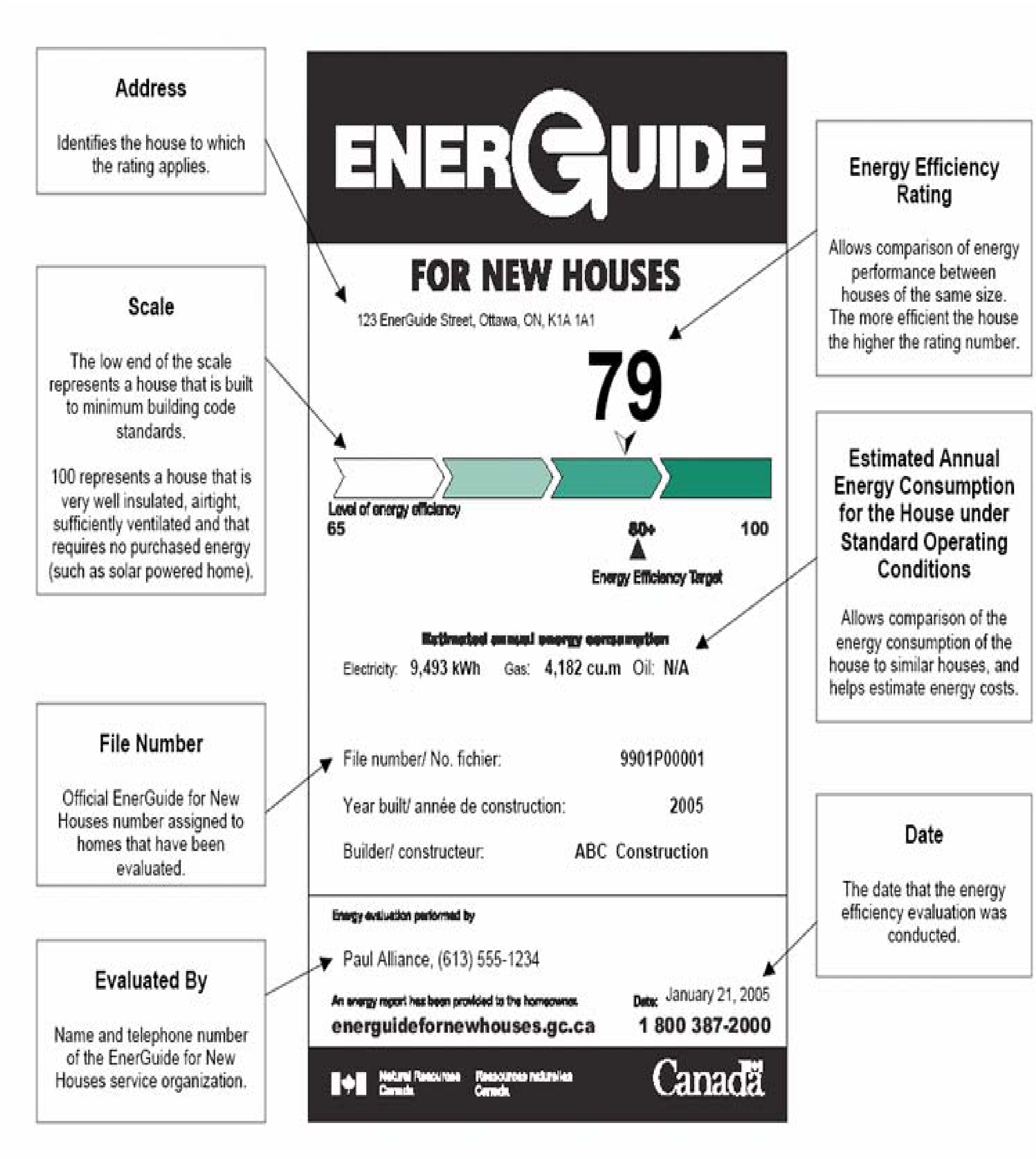




"The line on the QQ plot shows expected values for a normal distribution- the closer the values to the line, the closer the distribution is to normal" (p207, Mitchell, 2005). The plot shows that the number of participating households' values is not normally distributed.

Appendix E

EnerGuide: Energy Efficiency Rating



(City Green Solution, 2008)

Appendix F

Summary of Statistic of Dependent Variable

Based on information collect by REEP office from 1999 to 2006, the average of number of participating households in the study area was around 10 households, and the mean participation rate based on CDA level was 6%. The minimum number of participating households was 0, and the maximum number of participating households was 66. While, the minimum participation rate was 0% and maximum was 51.56%.

Descriptive Statistics –Number of Participating Households - Based on CDA Level

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
# of Participating Households	694	0	66	10.26	8.595	73.879

Descriptive Statistics – REEP Participation Rate Based on CDA Level

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Participation Rate	694	0	51.56	6.0045	5.01213	25.121

Appendix G

Parameter Estimates Comparison: Weighted Dependent Variable vs. Unweighted Dependent Variable

Analysis of Parameter Estimates (Weighted dependent variable by weight_M)

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	**Sig.**
(Intercept)	**.263**	.2727	-.271	.798	.933	1	**.334**
AVHHINC	**3.188E-6**	1.2402E-6	7.575E-7	5.619E-6	6.609	1	**.010**
APPR_DWENO	**.001**	9.8424E-5	.001	.001	118.430	1	**.000**
PER_HIGH_E	**.029**	.0030	.024	.035	98.230	1	**.000**
EMPLOY_RTE	**.008**	.0036	.001	.015	5.662	1	**.017**
OWNED_PERC	**.005**	.0015	.002	.008	11.927	1	**.001**
PEROVER65	**.009**	.0042	.001	.017	4.428	1	**.035**
PERBF1970	**.006**	.0009	.004	.007	41.700	1	**.000**
(Scale)	4.980[a]						

Dependent Variable: T_HH_IN_PA * Weighted by Weight_M (Weight_M= number of eligible households in each CDA/mean of eligible households in study area)

Model: (Intercept), PER_HIGH_E, AVHHINC, EMPLOY_RTE, OWNED_PERC, PEROVER65, PERBF1970, PER_ELGIHO   a. Computed based on the Pearson chi-square.

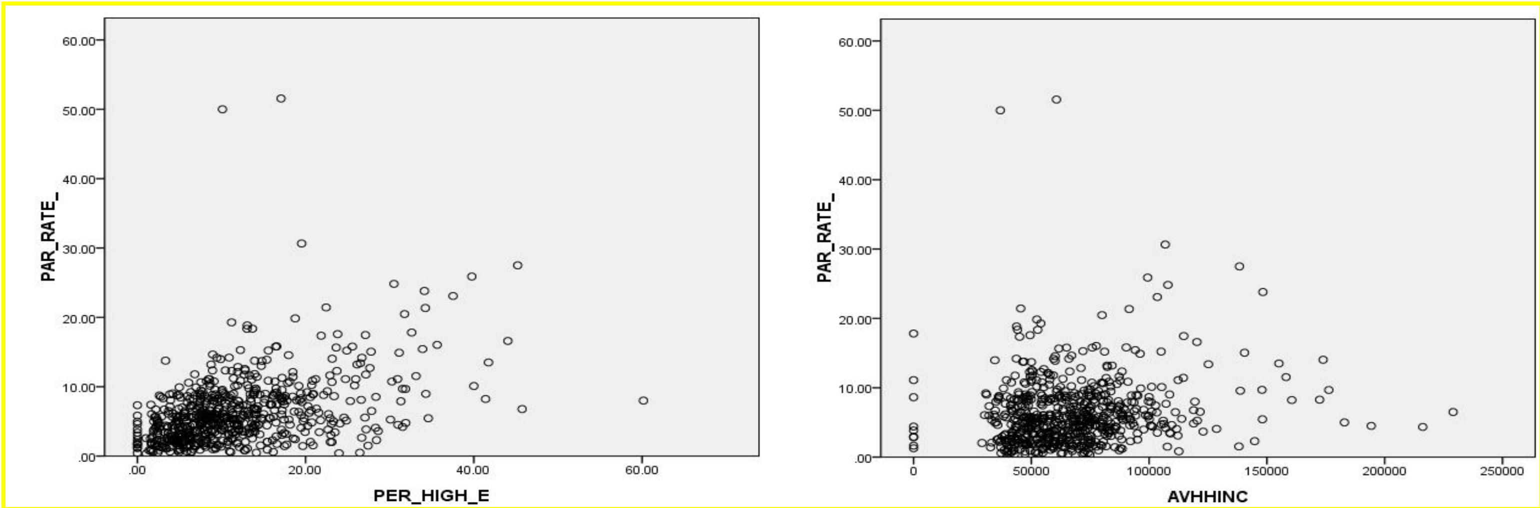
Analysis of Parameter Estimates (Unweighted dependent variable)

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-.248	.2702	-.778	.282	.842	1	.359
AVHHINC	2.343E-6	1.3488E-6	-3.007E-7	4.986E-6	3.017	1	.082
APPR_DWENO	.001	.0002	.001	.002	78.768	1	.000
PER_HIGH_E	.030	.0031	.024	.036	92.054	1	.000
EMPLOY_RTE	.010	.0035	.003	.017	7.548	1	.006
OWNED_PERC	.009	.0014	.006	.012	37.411	1	.000
PEROVER65	.005	.0041	-.003	.013	1.695	1	.193
PERBF1970	.008	.0008	.006	.009	85.638	1	.000
(Scale)	4.726[a]						

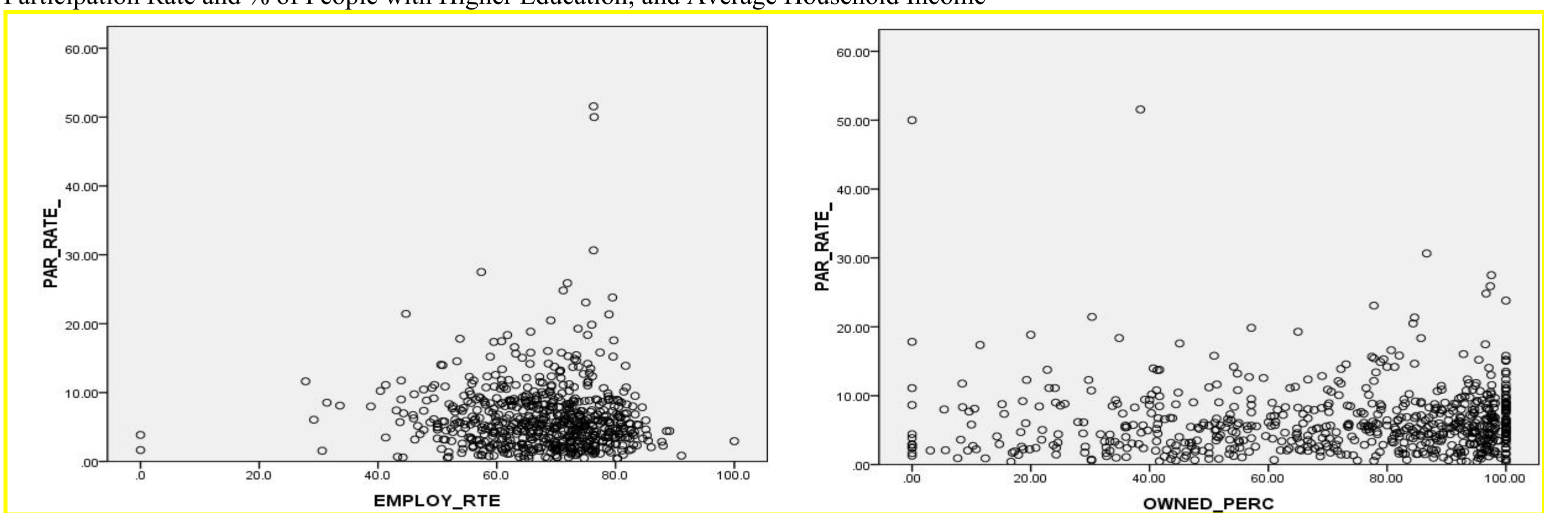
Dependent Variable: T_HH_IN_PA *unweighted dependent variable.

Appendix H

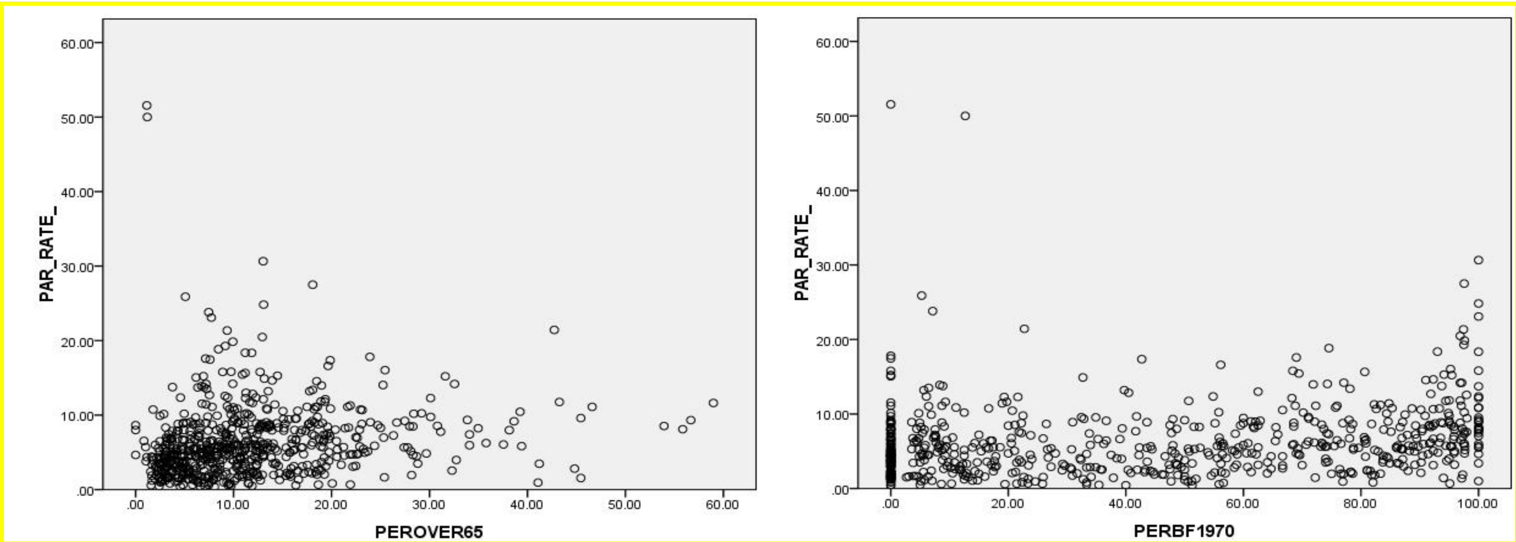
Scatter Plot



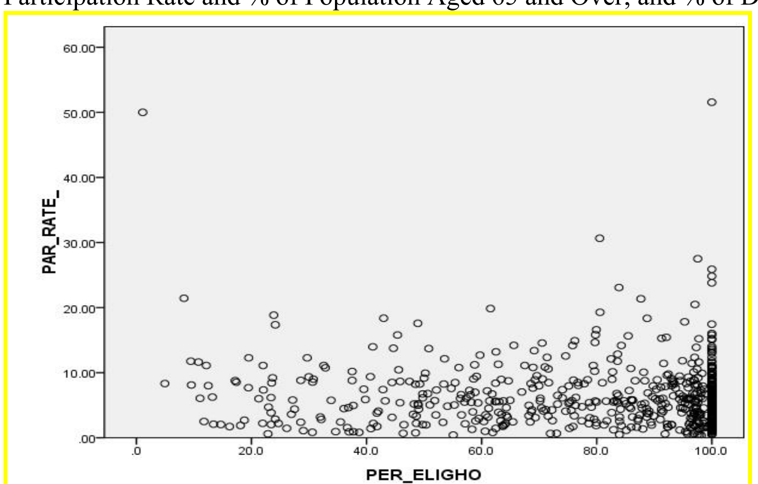
Participation Rate and % of People with Higher Education; and Average Household Income



Participation Rate and Employment Rate; and % of Owner Occupied Dwelling



Participation Rate and % of Population Aged 65 and Over; and % of Dwellings Built before 1970



Participation Rate and % of Eligible Households