uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Noura Algwaiz

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Electronic Business Technologies)

GRADE / DEGREE

Telfer School of Management

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Modeling Internet Penetration in Canada

TITRE DE LA THÈSE / TITLE OF THESIS

Bijan Raahemi

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

Morad Benyoucef

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Craig Kuziemsky

David Wright

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Modeling Internet Penetration in Canada

By

Noura Algwaiz

Thesis Submitted to the

Faculty of Graduate and Postdoctoral Studies

In partial fulfillment of the requirements

For the degree

Master of Science in e-Business Technologies

School of Information Technology & Engineering- Telfer School of Management

Faculty of Graduate and Post Doctoral Studies

UNIVERSITY OF OTTAWA

Thesis Supervisor: Prof. Bijan Raahemi

Thesis Co-supervisor: Prof. Morad Benyoucef

Your file Votre référence
ISBN: 978-0-494-73813-9
Our file Notre référence
ISBN: 978-0-494-73813-9

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my gratitude to those who helped me conduct this research. First, I would like to thank Professor Bijan Raahemi, my supervisor, and Professor Morad Benyoucef, my co-supervisor. They guided me throughout this research by giving me the time to discuss each step and address my concerns. They were also very considerate and understanding on the personal level, which had a positive impact on my experience. I would like to thank Professor Dominique Ferrand, director of the e-Business program, for his course "Research Methodologies" which I found to be very helpful throughout the journey of conducting a research and writing a thesis. Professor Ferrand also took the time to read my proposal and provide constructive criticism to the direction of my research. I would like to thank Susan Mowers, the librarian who helped me obtain the data for this research. I would like to thank Dr. Jean-Michel Billette of Statistics Canada for answering my questions regarding the user guide for the CIUS 2007. I would like to thank Professor Pierre Brochu from the Department of Economics for answering my questions regarding the interpretation of a logistic regression model. Last but not least, I would like to thank my husband Khaled for his support which helped me overcome the difficulties and embrace the learning experience.

Abstract

Internet penetration is an important measure for a knowledge-based economy as it indicates how connected it is to the internet. It does not spread evenly across regions and societies, which results in digital divides. Despite being one of the most connected countries, Canada suffers from uneven penetration rates across the country. In this research, we study the socio-economic factors that influence Internet penetration in Canada. We found that the influence of rurality has decreased between 2005 and 2007, which suggests that initiatives that took place in those two years were effective. We also analyze the differences among the regions of Canada and found that the regions least influenced by the demographic variables are not necessarily the ones with the highest penetration rates. Therefore, a mere look at the penetration rates across the regions is not enough to assess the connectivity of the region and the digital divide within.

Chapter One: Introduction

Information and Communication Technology (ICT) is crucial for countries and societies to develop and compete in today's knowledge economy. It has the potential to help societies overcome social divisions by empowering individuals with knowledge and connectivity to government and public services. It also has the potential to increase social exclusions for those who do not have access to it. There are varying penetration rates among communities based on socio-economic variables, such as income, education and ethnicity. The unequal penetration rates have resulted in a digital divide; the divide between individuals in terms of their access to and use of technology. Industrialized countries are investing in ICT to ensure that their citizens are not left behind [1].

Internet penetration is defined as the ratio of number of Internet users to the total population of the country. This measure is an important indication of how accessible the information is to the citizens of the country. Canada has worked towards increasing internet penetration and benefiting the society from its advantages. With a penetration level of 73.2% in 2007, Canada is one of the leading countries in terms of internet penetration [25]. However, this image of Canada masks unequal internet adoption patterns within the country [3]. According to results for 2007, the digital divide still exists among groups of society based on income, education and age. In addition, individuals residing in rural areas are less likely to use the internet. There are also differences in penetration rates among the regions of Canada. British Colombia, Alberta and Ontario continue to score higher rates of penetration than the national level. On the other hand, Quebec and the Atlantic provinces continue to score lower rates than the national level [25].

Most of the research on the internet digital divide in Canada has focused at the national level, analyzing the differences between the urban and rural population, or between the different age

groups. To our knowledge, there are no studies that focus on analyzing the regional differences in internet penetration. In this research, we analyze the factors that influence internet penetration in Canada on the regional level. We analyze the influence of the factors on each region in order to assess the importance of each factor in each region.

McKeown and Noce (2007) built a logistic model to study the variables that influence internet use in Canada based on the Canadian Internet User Survey (CIUS) of 2005 [8]. The results provide a quantification of the influence of each variable on internet use in Canada. In this research, we built an updated model for 2007 and compared it with the results of the initial study. Then, we applied it to the different regions of Canada in order to compare the results among the regions. We also added a variable "Privacy Concerns" to the regional models as a proxy for comfort with using the internet. Novice and less skillful users tend to have more privacy concerns about using the internet [25].

1.1 Objectives

The first objective is to establish an econometric model to study the determinants of internet penetration in Canada and compare our results with the original model of 2005 by Mckeown and Noce. We revisit the study and apply the same model to the data available for 2007, then compared our results for 2007 with the original model. Our aim here is to see the change in the influence of each variable on internet use during these two years. Although two years might seem like a short period to see a difference, it is not the case in the rapidly changing ICT industry. The penetration rate of the internet has changed in Canada by 5% in two years [25]. Looking at the change in the model could help us understand where the 5% increase came from.

The second objective of this research is to analyze these factors for each region and provide a quantification of the relative importance of each factor on internet penetration per region. We

would like to know whether the effects of the variables are equal for all the regions, or whether they have varying degrees of importance across Canada. We also add an additional variable "Privacy Concerns" as a proxy for comfort with using the internet. This is only applied to the regional models for the sake of comparison.

1.2 Motivation

A future objective behind this research is to use the results of the analysis to estimate internet penetration based on the factors that influence internet penetration. The internet penetration is currently calculated based on the surveys conducted biannually in Canada. The use of a model to estimate internet penetration based on the influence of the socio-economic factors would eliminate the need for surveys to calculate this value.

1.3 Contributions

The influence of the socio-economic factors (such as age, income, locations, etc.) that impact internet penetration changes over time. Some factors' influence gradually decrease or increase over time. We have seen this with the gender factor in 2005, as the odds of internet use in Canada became higher for a female than for a male, although it was the opposite when the internet was first introduced in Canada [8]. ICT is a rapidly changing industry, therefore we believe that there is a need to monitor the influence of these factors in order to update and adjust programs and initiatives accordingly.

The results of this study provide an updated model of the socio-economic factors that influence internet penetration in Canada. We also analyze the influence of these factors per region and include an additional variable to the model "Privacy Concerns" which has not been included in

previous studies. The results of this study provide insight for the direction of future initiatives to promote internet adoption and related services such as e-government and e-commerce services.

1.4 Scope of the Research

In this research, we analyze the determinants of internet penetration in Canada for the year 2007 in six regions: (1) The Atlantic Provinces excluding Prince Edward Island (PEI), (2) Quebec, (3) Ontario, (4) Manitoba and Saskatchewan, (5) Alberta, (6) British Colombia (BC). We used the Canadian Internet User Survey (CIUS) 2007 as our data source. This is a large dataset which allows for a per region analysis. Yukon, Northwest Territories and Nunavut have been excluded from the survey and therefore have also been excluded from this research. PEI was also excluded from this research because the survey's design restricted us from analyzing the geographical location factor for PEI, one of the determinants of internet use in Canada. The unit of analysis is the individual.

1.5 Research Methodology

We study the influence of the socio-economic variables by applying a regression model to the CIUS 2007 dataset. This is a relatively large data set, 26588 cases for 2007, which is an advantage when applying regression analysis as it increases the robustness and reliability of the results. Due to the structure of the data, we apply logistic regression and examine the odds ratio of each variable in the model. The software we use is PASW, the latest version of the SPSS software. We build 8 models:

1- A validation model based on the CIUS 2005. We used this model to compare it with the original model and assess the reliability of our implementation. We used the same variables in the original model, except for the language variable which we were not able to obtain.

   Variables in this model: age, income, education, location, gender, and presence of children[1]

[1] For variable descriptions, check section 4.3 Data Preparation.

2- After validating our model, we build a model based on the CIUS 2007 to compare it with the original model and find out how the variables changed in two years. Here, we used the same variables from the original model except for the "presence of children" which was found to be not a significant variable in 2005, and the language variable which we were not able to obtain.

   Variables in this model: age, income, education, location, and gender

3- We use the CIUS 2007 to build six models, one for each region. Here, we include the variable Privacy Concerns which was not included in previous studies.

   Variables in this model: age, income, education, location, gender, and privacy concerns

1.6 Organization of the Thesis

In this chapter, we presented our objectives for this study. Chapter two discusses internet use in Canada and explains the Digital Divide and Internet Penetration which are primary concepts in our research. In chapter three, we provide descriptive analysis of our data. The Implementation of our study is explained in chapter four. In chapter five, we present the results of our study along with a discussion. Finally, we present our conclusions and future work in chapter six.

Chapter Two: Literature Survey

Since the internet started, Canada has been one of the most connected countries in the world. As of January 1st 2005, the ITU ranked Canada fifth in broadband network penetration worldwide [2]. Nonetheless, According to Middleton and Sorensen, Canada's image of being a country with high internet penetration rates "masks" the disparities in internet adoption [3]. Although internet availability and usage rates have increased, the digital divide between urban and rural areas did not change overtime despite policies addressing the issue [8]. Middleton and Sorensen also believe that the Canadian internet adoption was not researched enough. According to our review of the literature, we have found that the research in this area analyses and identifies the determinants of internet use across Canada. The disparities between individuals are researched based on the identified determinants of internet use. For example, the location of the individual is a determinant of internet use, and there is research on the disparities in internet access and usage between urban and rural individuals. Yet, there are disparities in penetration rates among the provinces and regions of Canada and we are not aware of any research on this matter. In this chapter, we discuss the Digital Divide and Internet Penetration. Then, we will focus on Canada by presenting some prior research and government initiatives.

2.1 The Digital Divide

When a technology is first released, it does not spread evenly across countries and regions. There are factors that influence the spread of the technology. These factors can be physical, such as the availability of technology infrastructure, social, such as the society's interest in the technology, political, such as government initiatives to promote or control the use of the technology, or economic, such as the cost of utilizing the technology. Different factors exist depending on the

technology and the society. For example the factors that influence internet penetration in North America are different than those in the Middle East.

The diffusion pattern of a technology is also affected by its relationship with other technologies in terms of how they influence the technology's functionality. For example, after the introduction of the internet, computers found a new usage and became necessary to access the internet. In other words, the introduction of the internet had an influence on the functionality of the computer and consequently an influence on its diffusion [5]. On the other hand, DVDs have replaced the functionality of VCRs and therefore caused the technology to be obsolete.

The uneven diffusion of technology has caused what is called today the Digital Divide. The Organization for Economic Co-operation and Development (OECD) defines the Digital Divide as "the gap between individuals, households, businesses and geographic areas at different socio-economic levels with regard to both their opportunities to access information and communication technologies (ICTs) and to their use of the Internet for a wide variety of activities". Many developing countries have computer and internet penetration rates that are equivalent to 1/100th of the rates in developed countries [4]. The gap between developed and developing countries is referred to as the global digital divide. There are also internal gaps within one country among different geographical regions or social classes [4].

Sciadas (2002) discusses the digital divide in ICT [5]. Divides in ICTs depend on the technology, the timing of its introduction and the variables of interest (income, education, age, etc). In general, income has a larger influence on the penetration of new ICTs, such as cell phones and the internet, rather than on established ICTs, such as televisions (see Graph 1). This is because growth rates are initially very high among high income groups, then the growth rates become higher for low income groups as the high income groups approach saturation and the technology spreads among the lower

income groups. This is when the divides start to narrow.

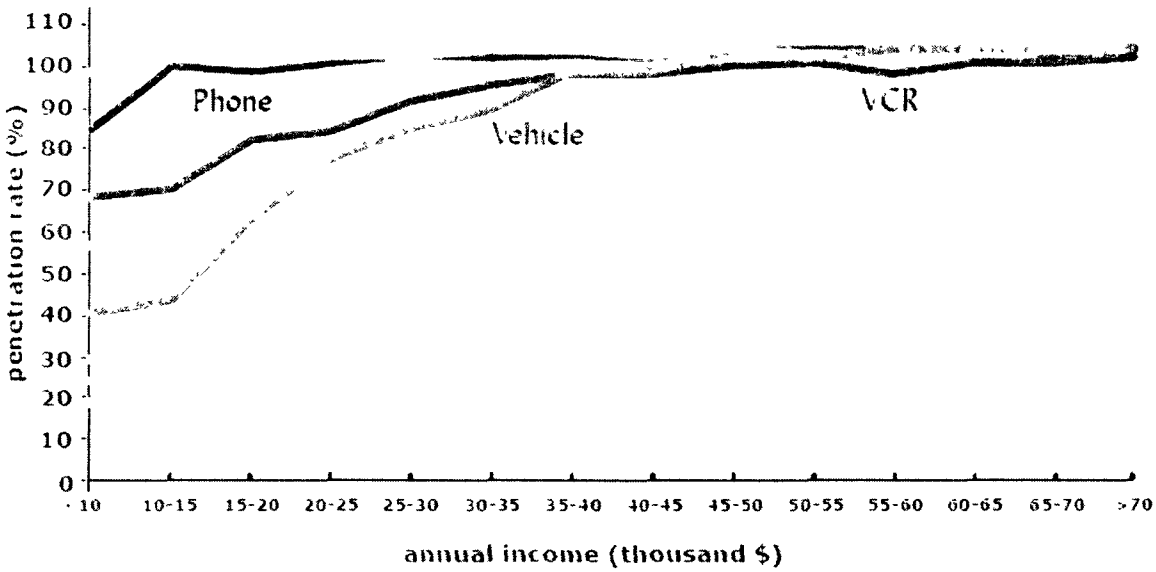


Figure 1: Household penetration of different technologies by income (2000) [5]

Research and initiatives have taken place to address and reduce the digital divide. The International Telecommunication Union (ITU) is the leading United Nations agency for information and communication technology issues. It has developed the ICT Opportunity Index, a tool to track the digital divide by using a list of indicators to measure the relative difference in ICT opportunity levels among economies and over time [6]. The ITU has identified ten indicators and categorized them to four sub-indices:

1) Networks

- Fixed telephone lines
- International Internet Bandwidth
- Mobile Subscribers

2) Skills

- School Enrolment
- Literacy

3) Uptake

- Computers
- Internet Users
- Households with TVs

4) Intensity

- Broadband Subscribers
- International Outgoing Tele-traffic

2.2 Internet Penetration

The digital divide is usually an issue for newly introduced technologies. As the technology matures, the gaps reduce and the penetration levels get closer to saturation. The internet is a relatively new technology suffering from uneven penetration rates. The percentage of Internet Penetration is defined as the number of internet users per 100 inhabitants (i.e. the number of internet users divided by the total population, then multiplied by 100.) Growth rates are now higher for developing countries than for developed countries. This is explained by the S-Curve. Penetration levels in developed countries are getting closer to saturation and therefore there is less room for growth, whereas the developing countries are in the middle of the curve which is a period of rapid acceleration (see Graph 2).

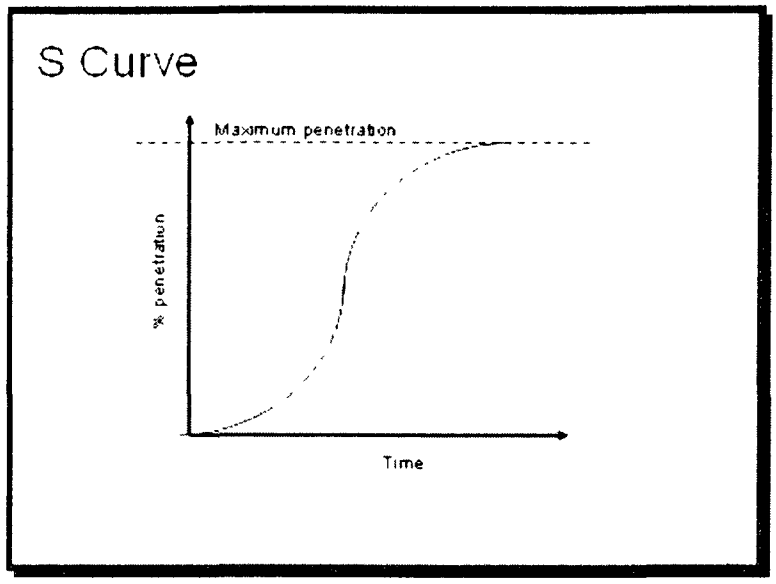


Figure 2: The S-Curve, Source: http://thebreakthrough.org/blog/s-curve.jpg

Identifying the determinants of Internet Penetration

Research is taking place to identify the determinants of internet use globally and regionally. After reviewing the research in this area, we have found that prior studies use regression analysis to determine the influence of predictor variables (also called independent variables) on the response variable (also called dependent variable) which is internet use in this case.

Chinn and Fairlie (2006) followed an econometric approach to analyze the determinants of computer and internet penetration rates [4]. A panel of 161 countries, which includes developed and developing countries, was examined over the period 1999-2001. According to this study, the global digital divide is mainly, but not entirely, caused by the different income levels. The variables included in this study are classified into five categories:

Economic Variables:

- Gross national income per capita
- years of schooling
- illiteracy rate
- trade openness

Demographic Variables:

- youth and aged dependency ratios:
    - population ages 0-14 (% of total)
    - population ages 65 and above (% of total)
- urbanization rate: urban population as a percentage of the total population

Infrastructure Indicators:

- telephone density: main telephone lines per 100 people
- electric power consumption (measured as kwh per capita)

Telecommunications Pricing Measures

- monthly telephone subscription charge
- cost of a 3 minute local call

Measures of institutional quality or policy:

- regulatory quality

The study uses regression analysis to analyze the relative importance of these variables in contributing to computer penetration and internet penetration on a global scale. The variables above are the independent variables in the models. They are analyzed first for the dependent variable

"Computer Penetration" which is measured as number of PCs per 100 people. Then, the same independent variables are analyzed for the dependent variable "Internet Penetration" which is measured as the number of internet users per 100 people.

The results of the study assure that income has the strongest effect on PC and internet penetration. Infrastructure indicators have a strong effect on the global digital divide. The region that is most affected by this variable is Sub-Saharan Africa. The quality of regulation is of great importance. Regions that are most affected by this variable are the Middle East and North Africa. These countries need to improve quality regulations in order to increase the uptake of the internet. Developed countries comprise a lower percentage of children and a higher percentage of elderly than developing countries. Since age has a negative effect on internet use, age distributions in the developed countries pose a disadvantage for these countries.

The results of this study suggest three directions to narrow the digital divide:

- public investment in human capital
- increasing the availability and quality of the telecommunications infrastructure
- improving regulatory quality

Statistics Canada

Statistics Canada is a primary resource for internet use in Canada [3]. The Household Internet Use Survey (HIUS) gathered information about access to the internet on the household level. This put a constraint on the research that can be done on internet use in Canada. The director of the Science, Innovation and Electronic Information Division at Statistics Canada notes: "While households can have an income, they cannot have an education or a sex" [3]. The HIUS was conducted from 1997 to 2003, then, it was replaced with the Canadian Internet User Survey (CIUS). The CIUS was

conducted in order to be more in line with international standards of ICT data collection set by the OECD. It gathers information about internet use on the individual level rather than the household level. The CIUS defines an internet user as "Someone who accessed the Internet from any location for personal, non-business reasons during a 12-month reference period" [25]. According to Statistics Canada, Quebec and the Atlantic Provinces continue to have lower internet penetration levels than the national rate, while the western provinces continue to have higher rates. Even when looking at other technologies, such as cell phones, western provinces have higher rates than Quebec and the Atlantic Provinces [7].

Table 1: Internet penetration in Canada and per province for 2005 and 2007

Geography	2005	2007
Canada	67.9	73.2
Newfoundland and Labrador	55.0	61.1
Prince Edward Island	60.6	69.3
Nova Scotia	67.1	69.4
New Brunswick	57.2	64.7
Quebec	62.4	69.0
Ontario	71.8	74.8
Manitoba	65.9	70.1
Saskatchewan	66.2	72.7
Alberta	70.6	77.4
British Columbia	69.3	77.6

Source: Statistics Canada, Table 358-0122 : Canadian Internet use survey, Internet use, by location of access, Canada, provinces and selected census metropolitan areas (CMAs), every 2 years (percent).

2.3 The Determinants of Internet Penetration in Canada

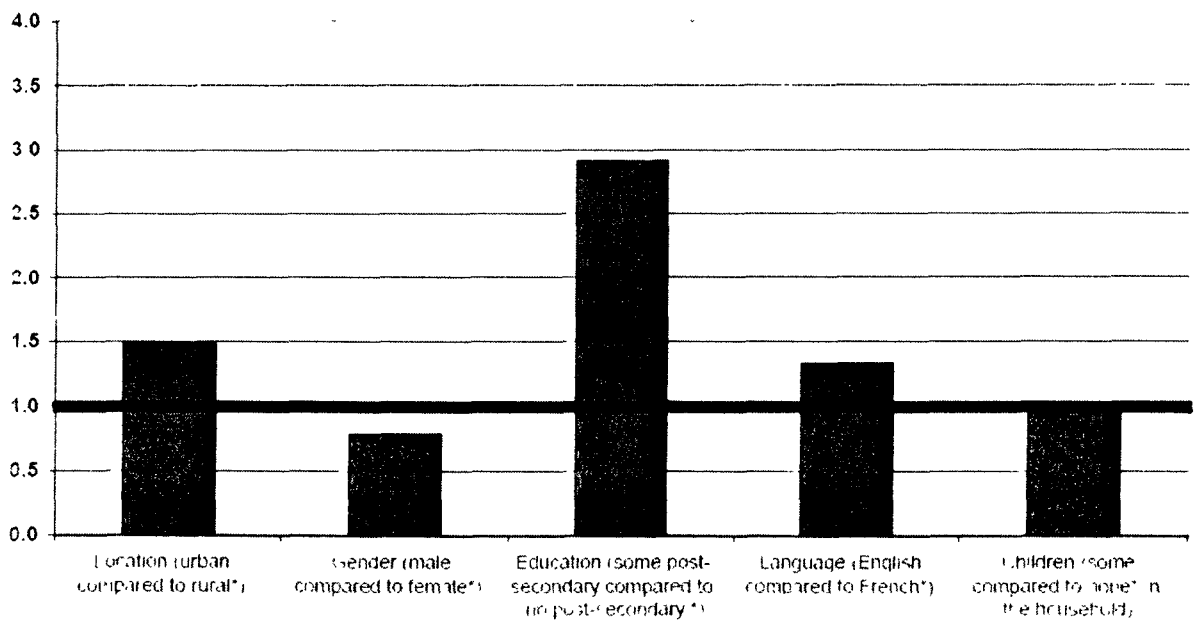
McKeown and Noce (2007) studied the effect of demographic variables on internet use in Canada [8] A stepwise logistic regression was applied to the following variables which have been identified to be the most influential variables on internet use in Canada

- age (continuous variable)
- household income (continuous variable)
- geographical location (binary variable, indicating whether the respondent resides in an urban or rural area)
- sex (binary variable, indicating whether the respondent is a male or female)
- education level (binary variable, indicating whether the respondent has obtained post-secondary education or not)
- Language (binary variable, indicating whether the respondent was interviewed in English or French. This is used as a proxy to indicate whether the respondent speaks English or French)
- Presence of Children (binary variable, indicating either the presence of children under 18 in the respondents' household or not)

Since the method of analysis used was a logistic regression, the results were reported in the form of odds ratios. The results indicate that all the variables entered in the regression have a statistically significant effect on internet use, except for the presence of children. This variable had a greater influence in the past but the results suggest that the influence of this variable has diminished over time Income remains a significant determinant of internet use but its significance is decreasing For a 1000$ increase in household income the odds of being an internet user increase by 1% The age of the respondent influences internet use such that every additional year results in approximately a 6% reduction in the odds of using the internet Education and location also remain significant

determinants of internet use, with the odd of using the internet being three times greater (e^{B} = 2.96)[2] for someone with at least some post-secondary education, and approximately one and a half times greater (e^{B} = 1.44) for someone living in an urban area. The odds of internet use were 1.3 times greater for someone interviewed in English.

An interesting finding is that being a male in Canada decreases the odds of accessing the internet by 20%. Although the percentage of internet use is equal for males and females, the odds of using the internet for a female are higher after controlling for the other variables. On an international level, there is a divide in internet penetration between males and females; being male has a positive influence on internet use, especially in developing countries. This gender gap was found to be narrower and almost diminished in some of the developed countries. Interestingly, this study indicates that the gender effect has swung to the opposite direction over time in Canada.



* reference group

Figure 3: Odds ratios for the determinants of Internet Penetration in Canada [9]

[2] e^{B} is a symbol of the Odds ratio, which is the indicator of the level of influence of an independent variable on a dependent variable in a logistic regression model.

The study concludes that influences on internet use are changing over time which emphasizes the importance of frequent research and up to date results. The results of the model suggest that education and location are the strongest determinants of internet use. Therefore, Canadian policy must continue to address both supply and demand by increasing the availability of infrastructure in rural areas and providing training to use the internet.

McKeown, Noce & Czerny (2007) study the effect of rurality on internet use in Canada [9]. Rurality is the extent of being rural. It is greater for individuals living in places with lower population densities and/or places further from an urban center. The gap in internet use between urban and rural areas is caused by socio-economic factors related to each geographical area. Populations in urban areas are generally younger and have higher levels of income and education. Another reason for this gap is the lack of internet availability, and more recently, the lack of broadband availability in rural areas.

Rural areas present a less attractive investment for telecommunication market forces due to the lower population concentration in these areas. Yet, some argue that the principal driver for providing ICT services to these areas should be the market forces. Others argue that these areas require further assistance from the government to be connected to the internet and participate in the information society.

Middleton and Sorensen (2005) analyzed Canadian households that are not internet users [3]. The study included "drop-outs"; those who used the internet more than a year before the survey was conducted and are no longer internet users. They distinguish between two types of non internet users: near and far. The near non-users are those who want to use the internet but do not have access, often because of high cost or lack of availability. The far non-users are those who are not interested in using the internet. According to Middleton & Sorensen, it is suggested that internet

access for the "near" users can be identified based on the socio-economic factors that determine internet use. On the other hand, internet access for the "far" users is determined based on interests and needs to access the internet; i.e. based on behavior. Distinguishing between these two types of non-users is called the Dual Digital Divide. The first divide occurs between those who use the internet and those who don't. The second Divide is among those who don't by classifying them as either near non-users (interested) or far non-users (non-interested).

The study analyzed five years of the HIUS to study the households that do not use the internet. It has found that the adoption rate is influenced by age, education, income and family type. The most frequently reported reasons for not using the internet were either that the household doesn't need to use the internet or is not interested. The second most frequently reported reason was cost, followed by lack of time, then lack of skills. The study also analyzed "Dropout" households, households that used the internet more than a year before the survey was conducted. The most reported reason was a lack of access to a PC, followed by a lack of need, then the cost of internet access being too high. The trend for reporting the first two barriers was decreasing, whereas the trend for reporting the third reason was steady. The findings of this study demonstrate the Dual Digital Divide; some households are near non-users because the cost of the internet limited their access to the internet, and some households are far non-users because a lack of interest limited their access to the internet.

According to the authors, applying logistic regression to the variables that determine internet use for households in Canada (age of the head of the household, education level of the head of the household, the household income, the family type, and the labour force status of the head of the household) for the HIUS predicts the households that use the internet. However, using these variables as determinants of non-use of the internet is not effective. To illustrate this, a logistic regression was applied to the HIUS 2003 and accurately classified 91.6% of households using the internet, but accurately classified only 67.3% of households that do not use the internet. The reason

is that these variables can identify the near non-users but not the far non users.

The study concludes that it is extremely important for all Canadians to have an equal opportunity to access the internet and participate in the information society, but also it is important to give them the option not to participate if they prefer. This implies that the focus should be on reducing the digital divide among the near non-users rather than the far non-users who have the opportunity to use the internet but choose not to use it. The study also believes that the "Connecting Canadians" program, a federal government initiative to provide equal access to the internet, has failed to connect many Canadian households, precisely those who are older and/or less advantaged among the society. We will discuss government initiatives including the Connecting Canadians program in the next section.

Belanger and Carter (2009) also distinguish between two kinds of divides: an access divide and a skills divide [10]. The access divide refers to the lack of physical access to the internet. Ethnicity, income, education and age are significant predictors of the access divide. On the other hand, the skills divide is the lack of the necessary skills to effectively use the internet.

Cullen (2001) categorized barriers to internet use into four key issues [11]:

- **Physical access:** this is determined through the availability and robustness of the telecommunication infrastructure, the cost of accessing the internet, and support of access for the disabled.
- **Lack of ICT skills:** this includes computing skills and literacy skills.
- **Attitudinal Barriers:** this includes negative perceptions towards the internet, such as concern over security, concern over the suitability of the internet's content for children, the impression that the internet is difficult to use, and the impression that the internet is for certain groups of the society, such as males or youth.

- **Content:** lack of relative content to the user presents a barrier to internet use

According to Cullen, once the physical access is available and the skills are developed, it is easier to change the attitudes and develop relative content

From the studies above, we conclude that efforts to increase internet adoption and reduce the digital divide should address supply and demand Addressing supply is by increasing internet availability and accessibility Addressing demand is by improving ICT skills and investing in human capital.

The Canadian Internet Project

The Canadian Internet Project (CIP) is an ongoing longitudinal study of the trends of Canadians in internet use and attitudes towards it [12]. CIP is a member of the World Internet Project (WIP), an academic consortium of research centers in 28 countries around the world. The WIP publishes an annual international report comparing data about internet use from the member counties The CIP conducts a survey every three years, the first was in 2004

According to the CIP, the internet, technology and media influence the attitudes and daily lives of Canadians like no other element of society whether political, cultural, economic or social Consequently, it is important that the industry, policy makers, service producers and the public are well informed about the Canadian trend in using the internet to assess and respond to its impact. D. Jeffrey Cole, founder of the WIP, believes that it is important to study the internet in its early years before it matures in order to understand its impact He believes that the opportunity of studying the technology was missed for television and should not be missed for the internet [12]

In 2007, 3,150 telephone interviews were conducted with Canadian internet users and non-users 12

years of age and older from the ten Canadian provinces. Consistent with Statistics Canada, the results of the survey reveal that penetration rates in Quebec and the Atlantic Provinces are lagging behind the rest of Canada. The lower penetration rates in the Atlantic Provinces are explained by a lack of infrastructure. On the other hand, Quebec is lagging behind for two reasons. First, language presents a barrier to French speaking Canadians in Quebec due to the dominance of the English language on the internet. Second, the large rural areas lead to a lack of infrastructure in rural Quebec. However, the Prairie Provinces face the same geographical barrier as Quebec, yet have higher internet penetration levels. In fact, according to the CIP survey, the Prairie Provinces experienced the most dramatic growth in internet penetration from 2004 to 2007. Quebecers in rural areas are much less likely to be online than residents of rural areas in the other provinces. The results of the CIP imply that Quebec is the province most affected by the geographical variable (urban vs. rural).

Table 2: Percentage of penetration rate per region

Region	2004	2007
Atlantic Provinces	67	69
Quebec	68	64
Ontario	75	81
Prairie Provinces	69	81
Alberta	76	83
BC	74	82

Source: Canadian Internet Project 2007 [12]

Penetration rates across the regions of Canada are different from what was reported by Statistics Canada. This could be due to the different size of the sample in each survey. Also, the definition of

an internet user is different in each case. The CIP defines an internet user as a respondent who has used the internet during the three months prior to the survey, whereas for Statistics Canada the duration is 12 months. The CIP does not specify the purpose of usage as Statistics Canada does.

Table 3: Percentage of individuals with more than one computer at home (2007)

Region	Percentage with more than one Computer at home
Atlantic Provinces	31
Quebec	36
Ontario	45
Prairie Provinces	35
Alberta	46
British Colombia	47

Source: Canadian Internet Project 2007 [12]

Residents of BC, Alberta and Ontario are more likely to have more than one computer at home.

We believe that the CIP's results on the regions of Canada are not very reliable compared to those of Statistics Canada because the dataset for the study was comprised of 3,150 cases. Breaking this number down to the 6 regions would result in a small number of cases for each region, which decreases the reliability of the results.

2.4 Government Initiatives in Canada

There has been a debate as to whether ICT services should be supported by the government or left to the market forces. We believe that leaving the issue to the market forces would widen the digital

gap between urban and rural areas simply because rural areas are not an attractive investment compared to urban areas. Frieden (2005) supports our argument suggesting that government involvement is necessary for successful ICT incubation [2]. The study derives lessons from successful broadband development strategies in Canada, Japan and Korea. They found that successful strategies have focused on the following areas:

- Developing a vision and strategy
- Promoting digital literacy
- Investment in infrastructure
- Offering e-government services

Below, we present federal and provincial initiatives to increase ICT adoption in Canada:

1- Federal Initiatives

At the federal level, Industry Canada is in the lead in promoting the internet and broadband in Canada [13]. Federal initiatives were united under one program called "Connecting Canadians" which covers a wide range of issues organized around six pillars:

1- Canada On-line

2- Smart Communities

3- Canadian Content On-line

4- e-Commerce

5- e-Government

6- Connecting Canada to the World

One of the "Connecting Canadians" initiatives is the Community Access Program (CAP) that aims to increase the availability of affordable public internet access, such as in schools and libraries, and

the skills necessary to use the internet [14] Therefore, it addresses the physical availability of the internet and the lack of skills to use it

Rideout (2000) analyses the CAP to assess whether this initiative is bridging or aggravating the digital divide [15] The program provides funding to non-profit organizations such as schools, libraries, community centers, and municipal and territorial governments. The average federal funding is 20,000$ per project and covers the cost of computers, ISPs and training

According to Rideout, the CAP lifts income and social barriers to internet access. However, the program's focus on the technical access has excluded human social needs. Another problem is the need for continuous funding of the program to prolong its life. The CAP was charging users in order for it to survive. This is contradictory to the program's purpose since it was initially created to provide affordable public internet access to those who are less fortunate.

2- Provincial Initiatives

In addition to initiatives supported by the federal government, the provincial governments worked towards increasing connectivity within their jurisdictions. These initiatives launch or support programs to increase the availability of broadband in the province In some cases the focus is on the rural communities, such as initiatives in Ontario. In other cases, it is a provincial wide program We list some initiatives here:

- **Alberta:** The government of Alberta partnered with Axia SuperNet and Bell Canada to launch Alberta SuperNet, a network that provides broadband to the province including urban, rural and remote communities [16].
- **British Colombia:** Network BC was launched in British Columbia in 2005 to provide

broadband to rural and remote areas [17].

- **Saskatchewan:** The government of Saskatchewan partnered with SaskTel to form CommunityNet, a province wide broadband network [18].
- **Manitoba:** Between 2000 and 2008, the Canada-Manitoba Infrastructure Program funded projects such as "Rural and Northern Telecommunications Infrastructure" to provide ICT infrastructure rural areas, and "High-Speed Internet Access for Public Institutions" to support public facilities such as schools and libraries [19].
- **Ontario:** The government of Ontario supports several programs including Rural Economic Development "RED" and Connect Ontario: Broadband Regional Access "CORBA".
- **New Brunswick:** New Brunswick's eNB targets four components: e-Government, e-Learning, e-Business and e-Infrastructure [20]. Also, in 2008 the government of New Brunswick partnered with Barrett Xplore, a private internet service provider, to provide broadband to rural populations in the province [21].
- **Quebec:** The government of Quebec launched the Villages Branchés du Québec to increase broadband availability in Quebec schools and municipalities [22]. Quebec also initiated a program to connect families to the internet in order to reduce the digital divide caused by income disparities [23]. Families eligible for the program are those receiving a family allowance from "Régie des rentes du Québec", a program that contributes to the financial security of families and individuals [24]. The program enables these families to connect to the internet at a reduced cost and the option of obtaining a multimedia computer if desired. The service has a limited duration of two years.

Chapter Three: Data Analysis

Internet penetration is influenced by a range of variables. In Canada, the variables with the highest influence on internet use are income, age, education, geographical location, language, and gender. We used these variables to build the models for our study. Privacy concerns is a variable we added to the regional models to compare the effect of concerns over privacy in each region. We did not add Presence of Children since it was found to be not significant anymore [8]. Language was also not included because we were not able to obtain the data. In chapter two, we found that initiatives to increase internet adoption should address supply and demand. Therefore, we categorized our variables to two categories:

(1) Variables that indicate the need for supply: Income and location. These variables indicate the need to increase supply of the internet because they are associated with availability and accessibility to the internet.

(2) Variables that indicate the need for demand: age, education and privacy concerns. These variables indicate the need to increase demand for the internet because they are associated with the user's skills which are necessary to become an internet user.

In this chapter, we present the descriptive statistics for each region from the CIUS 2007 survey, our main data source for this study. In order to be able to interpret the results of our models, we need to study our independent variables in each region. We have used PASW, the software for our analysis, to create cross tabulations for each variable across the 6 regions. The tables in this chapter are all a result of cross-tabulations using the PASW software, except for table 4 (The percentage of Internet Users by community size).

3.1 Geographical Location

The geographical location variable (urban vs. rural area) is a determinant of internet penetration globally as well as in Canada [4]. Canada is challenged by its geographical nature which has led to a concentration of the population on the south border of Canada.

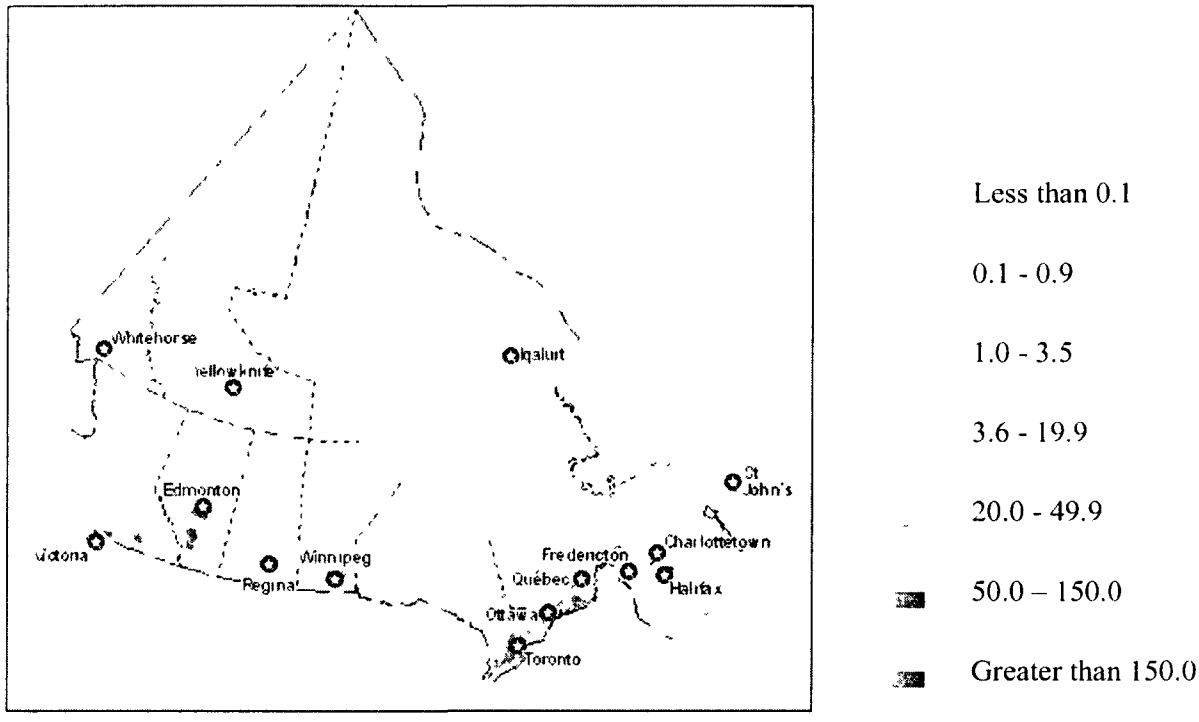


Figure 4: Concentration of the Population in Canada, Source: Natural Resources Canada

The divide between urban and rural areas is caused by a combination of socio-economic variables and the lack of internet or broadband availability. Urban areas present a younger population and higher levels of income and education [25]. Therefore, urban areas are a more attractive market to internet service providers when compared to rural and remote areas. This has resulted in a lack of internet availability, and more recently a lack of broadband availability, in rural areas. Table 4 demonstrates how the percentage of internet users decreases with the decrease of community size.

Table 4: The percentage of Internet Users by community size, 2005

Community size	Percent
CMA[3] > 1 million persons	73
CMA 250,000 to 1 million persons	71
CMA 100,000 to 249,999 persons	68
CA[4] (10,000 to 99,999 persons)	65
Rural and small town, strong to moderate MIZ[5]	56
Rural and small town, weak to no MIZ	57

Source: [9]

Rural areas are sparsely populated lands lying outside of census metropolitan areas (CMA) or census agglomeration areas (CA). They include small towns, villages and other populated places with less than 1,000 people, including remote areas [9].

We believe that this is a variable worthy of in-depth analysis and examination because there is a lot that governments can do to tackle the geographical digital divide and provide rural and remote areas with internet access. On the other hand, governments are limited in what they can do about the digital divide in other aspects such as age and gender. Also, the internet removes the distance barrier and presents connectivity opportunities to rural populations such as distant learning.

In chapter three, we discussed some of the government initiatives to provide broadband in rural areas. According to Noce and McKeown (2007), although penetration rates are increasing, the

[3] A census metropolitan area (CMA) has an urban core population of at least 100,000 (Statistics Canada)
[4] A census agglomeration (CA) has an urban core population between 10,000 and 100,000 (Statistics Canada)
[5] Metropolitan Influence Zone

divide between urban and rural areas has neither narrowed nor widened overtime despite government policies aimed at increasing the availability of broadband in rural areas and reducing the digital gap. In this study, we aim to find out which provinces are more affected by the location variable. To do this, we need to first consider the percentage of the rural population in each province. Table 5 lists the percentage of urban and rural population in each region. Ontario and BC have a considerably lower rural population compared to the other regions. The Atlantic Provinces and Manitoba/Saskatchewan have a higher rate than the rest of the regions. We need to consider this difference when interpreting the results of our model.

Table 5: The Percentage of Urban and rural population in each region.

Region of Respondent	Urban	Rural
Atlantic Provinces	57.2	42.8
Québec	75	25
Ontario	84.7	15.3
Manitoba/Saskatchewan	61.8	38.2
Alberta	72.8	27.2
British Columbia	78.8	21.2

Broadband Access and Availability

Since the location divide is caused by a combination of socio-economic variables and broadband availability, we looked into the availability of Broadband in each region. Table 6 lists the percentage of broadband access and availability per region. Broadband access is based on the number of respondents who indicated cable as their method of access or that they access the internet at home using broadband. Broadband availability is based on the respondents we identified for

broadband access plus those who indicated that broadband is available in their area but do not use it to access the internet. Table 6 indicates that BC has the highest level broadband penetration, followed by Alberta. The Atlantic Provinces have the lowest broadband access.

Table 6: Broadband access and availability for each region

Region of Respondent	Broadband Access	Broadband Availability
Atlantic Provinces	53.4	57.2
Québec	55.2	62.3
Ontario	63.0	67.9
Manitoba/Saskatchewan	59.0	63.0
Alberta	66.9	71.3
British Columbia	69.9	73.3

3.2 Age

Internet penetration rates decrease with the increase of age. However, this is a cohort rather than an age effect. Internet users do not stop using the internet as they age, but the younger generation tends to adopt internet use more than the older generation. Therefore, it is expected that the influence of age is gradually fading as the younger population of today grows older. The age of the respondent is recorded in five categories: 16 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 64, and 65 +. Since the median age of the population in Canada is 39.5[6], we divided the six categories to two breaking them in the middle at the age of 45. Table 7 indicates that Alberta has the youngest population of the regions whereas the Atlantic region has the oldest population.

Table 7: Percentage of respondents 16 to 44 years of age vs. 45+ years

Region of Respondent	16 to 44	45+
Atlantic Provinces	47.7	52.3
Québec	49.9	50.2
Ontario	52.3	47.6
Manitoba/Saskatchewan	50.3	49.8
Alberta	56.2	43.7
British Columbia	50	50

3.3 Income

Generally, the influence of income on new technologies, such as cell phones and the internet, is stronger than its influence on older more established technologies, such as telephones and televisions. When a technology is first introduced, the high income households are usually the majority of the adopters. As the technology matures and the high income households get closer to saturation, the adoption rate becomes higher for lower income households. This phenomenon could be explained by the decreasing relative cost of the technology as it matures. This has been observed for the internet, high income households are approaching saturation and the adoption rate for the internet is now faster for the middle income households. Also, the relative cost of access, bandwidth, and computer equipment is decreasing [5].

The income variable in the CIUS 2007 is divided into five categories: <=24k, 24k to 40k, 40k to 60k, 60k to 95k, 95k+. Since the median income in Canada is 63,600$[7], we used the 60k as a

[6] Source: Statistics Canada 2006 Community Profiles http://www12.statcan.gc.ca/census-recensement/2006/dp-pd/prof/92-591/index.cfm?Lang=E accessed October, 2009.
[7] Source: Statistics Canada Census 2006 http://www40.statcan.gc.ca/l01/cst01/famil107a-eng.htm

breakpoint and reported the percentages below 60k and more than or equal to 60k. We found considerable differences in income levels among the six regions. Alberta has the highest income, followed by Ontario. The Atlantic Provinces and Quebec have the lowest income levels.

Table 8: Percentage of respondents with incomes below 60k vs. incomes above or equal to 60k

Region of Respondent	<60,000	>=60,000
Atlantic Provinces	63.8	36.2
Québec	61.8	38.2
Ontario	46.2	53.8
Manitoba/Saskatchewan	55.5	44.5
Alberta	43.8	56.2
British Columbia	50	50

3.4 Education

Educational attainment is associated with using the internet and being a more sophisticated user. Those who have higher education levels tend to have the ICT skills needed to use the internet [11]. According to table 9, Manitoba and Saskatchewan have the lowest percentage of post secondary attainment.

3.5 Gender

Men tend to use the internet more than women, especially in developing and less developed countries. In some of the developed countries, the gap has diminished and in some cases has swung to the other direction. This is the case for Canada, McKeown and Noce (2007) found that the odds of using the internet for females are now greater than for males. According to Table 10, there are no

differences in the percentage of male versus female population among the regions worthy of consideration. The difference between the highest and lowest percentages is merely 2%.

Table 9: Percentage of respondents with post secondary education and without

Region of Respondent	High School or less	Post-Secondary Education
Atlantic Provinces	43.2	56.8
Québec	37.6	62.4
Ontario	39.8	60.2
Manitoba/Saskatchewan	46.7	53.3
Alberta	42.6	57.4
British Columbia	37.7	62.3

Table 10: The percentage of male vs. female population per region

Region of Respondent	Female	Male
Atlantic Provinces	51.7	48.3
Québec	51.1	48.9
Ontario	50.9	49.1
Manitoba/Saskatchewan	50.6	49.4
Alberta	49.7	50.3
British Columbia	50.8	49.2

3.6 Privacy Concerns

The attitude barrier refers to the negative attitude towards the internet, which includes a lack of

interest, the perception that the internet is too difficult, and concerns over security [11]. More experienced and sophisticated users tend to have fewer concerns over privacy and security on the internet [25].

The literature did not identify privacy concerns as a factor influencing internet use in Canada. However, we added this variable to our regional models as a proxy for comfort with using the internet. This variable will enable us to find out if privacy concerns have decreased the odds of using the internet or not for each region. The CIUS asks all respondents about their concerns over privacy on the internet, such as people finding out what websites they visited or others reading their emails. The answers are categorized to: (a) not concerned at all, (b) concerned, (c) very concerned. Bellow is the percentage of respondents for each category per region. Table 11 indicates that there are no considerable differences among the regions.

Table 11: The level of concern over privacy on the internet

Region of Respondent	Not at all Concerned	Concerned	Very concerned
Atlantic Provinces	26.7	34.9	38.4
Québec	24.9	36.6	38.5
Ontario	25.2	35.0	39.8
Manitoba/Saskatchewan	25.5	37.2	37.3
Alberta	24.5	36.9	38.7

Chapter Four: Implementation

4.1 Methodology: Logistic Regression

This is a study with a cross-sectional design. As we have illustrated in Chapter 2, prior research uses Regression Analysis to study the influence of variables on internet uptake. There are different types of regression analysis, the choice of analysis is based on the format of the data. Our dependent variable, internet use, is a dichotomous (binary) variable taking one of two values: either the respondent is an internet user or not For this kind of dependent variable, the applicable choice of regression is the binary logistic regression [30]. The software we are using for our analysis is PASW which is the latest version of SPSS.

According to Middleton and Sorensen (2005), logistic regression accurately predicts internet users using socio-economic factors. A logistic regression model was applied to the HIUS 2003 survey and accurately classified 91 6% of the households that use the internet. For non-users, logistic regression can accurately classify the near non-users but not the far non-users Therefore, it is not a reliable approach for classifying the non-users. McKeown and Noce also used logistic regression to model the determinants of internet use based on the CIUS 2005. Since we would like to compare our model with the original model, we will follow the same format of the variables in order to have comparable results.

The independent variables will be entered into the regression sequentially The variable that has a significant influence on the dependent variable and improves the model's fit stays in the model Variables that are not of significant influence do not enter the model This type of model building is

called stepwise and the significance level for keeping a variable in the model is 5% (p-value <= 0.05).

The Generic Logistic Regression Model:

$$Y = 1 / (1 + e^{-z})$$

where Y is the dependent variable, e is the base of the natural logarithms, approximately 2.718, and z is the linear combination:

$$z = \beta_0 + \beta_1(x_1) + ... + \beta_n (x_n)$$

Where x_1 to x_n are the independent variables under investigation and n is the number of independent variables.

Interpretation of the Odds Ratios

For Logistic Regression, the results are usually analyzed in the form of the odds ratio, symbol e^{β} [30]. The odds ratio represents the probability of an event occurring divided by the probability of the event not occurring, controlling for all the other variables. The odds ratio is interpreted depending on its value:

- $e^{\beta} < 1$, the variable decreases the odds of using the internet
- $e^{\beta} = 1$, the variable has no influence on using the internet
- $e^{\beta} > 1$, the variable increases the odds of using the internet

The odds ratio is interpreted differently when it is associated with a continuous variable than with a binary variable. For continuous variables, the odds ratio indicates the odds associated with a one unit change of the variable. For example, 0.96 odds ratio for age indicates a 4% decrease of the odds of using the internet for every additional year of age. For binary variables, the odds ratio indicates the odds that the respondent is an internet user when the variable is 1. For example, the variable gender takes the value 1 for a male respondent and 0 for a female. If the odds ratio for gender is equal to 2, then the odds of using the internet for a male respondent are 2 times greater than the odds of a female using the internet.

4.2 Data Source: Canadian Internet User Survey

The data we have used to conduct this research is from the Canadian Internet User Survey (CIUS). CIUS is a survey that measures the extent and scope to which individual Canadians use the internet. This is measured by the intensity of use, the frequency of use and the speed of the internet connection. The survey is conducted by Statistics Canada on a biannual basis. It has been conducted twice since it started, in 2005 for the first time and later in 2007. The 2007 survey has a total of 26,588 cases, this is a relatively large sample which allows us to conduct a per region analysis and achieve reliable results. The region with the smallest number of respondents is Alberta, which has 2,473 respondents.

The 2007 survey was modified to reflect feedback from respondents and interviewers from 2005 and to reflect the changes in internet use. The target population has also changed, it covered 18 years of age and older for 2005 and 16 years of age and older for 2007. CIUS replaces the Household Internet Use Survey (HIUS); a survey that was conducted on an annual basis from 1997 to 2003 to measure the household internet use. In order to be consistent with international standards

for internet statistics, the HIUS was replaced with the CIUS. This change enables international figures to be compared and studied in-depth [26].

The Ontario Data Documentation, Extraction service and Infrastructure (odesi) is a web based data exploration, extraction, and analysis tool. It provides the microdata file for the CIUS 2005 & 2007 to authorized users, including university students. This was the data resource for our study [27].

Scope of the Survey

Excluded from the survey were residents of the Yukon, Northwest Territories and Nunavut, Inmates of Institutions, Persons living on Indian Reserves, and Full time members of the Canadian Forces.

Deleted Cases

Prince Edward Island

The location of the respondent, whether he resides in an urban or rural area, enables us to study the effect of the location on internet use. For Prince Edward Island (PEI), the survey does not indicate whether the respondent resides in an urban or rural area because the PEI sample is very small and the confidentiality of the respondent might be lost if the survey indicated the location of the respondent. Since the location of the PEI respondent is not indicated, we were not able to run the model with the location variable on PEI. Consequently, PEI was excluded from the analysis.

Missing Values

After deleting cases belonging to either PEI or a multifamily household or both, a total of 4 cases were missing from each survey. For the regional models, we added the variable, privacy concerns. For this variable, there were some missing values either because the respondent doesn't know or

refused to answer or the answer was not stated. The cases with the missing values were deleted, which were a total of 1183 cases.

Table 12: Sample size per data set

Data Set	Sample size
CIUS 2005	28434
CIUS 2007	25882
Atlantic Provinces	3677
Quebec	4863
Ontario	7226
Manitoba / Sascatchewan	3645
Alberta	2473
British Columbia	2815

Weights

The Microdata User Guide for CIUS 2007 provides guidelines for analyzing the Microdata and publishing data derived from the survey [28]. In order to get valid results from our analysis, the survey weights must be used. Section 9.4, page 27 of the User Guide, specifies the method to be used when applying Logistic Regression to the data. The method calculates the average weight of each respondent, then rescales the weights by dividing them by the average weight. This results in new values for the weight variable, with a mean weight of 1. We applied this process to the CIUS complete dataset as well as for each region's dataset in order to build the model for each region.

4.3 Data Preparation

When applying logistic regression, independent variables should be either binary or continuous. Entering a categorical variable with more than two categories will result in more than one coefficient and more than one odds ratio for that variable. Therefore, we need to transform variables with more than two categories to either binary or continuous variables. We also need to be consistent with the variables used in the original model in order to have comparable results. In the original model, age and income were continuous variables and the rest were all binary.

The Dependent Variable

Internet Use (Y): The variable of interest in this study is internet use, a binary variable representing whether a respondent is an internet user or not. According to Statistics Canada, an internet user is someone who used the internet during the past twelve months from any location for personal non-business reasons [25]. In the CIUS survey, question PU_Q01 asks if a person used the internet during the past 12 months for personal non-business use. Assuming that Y is our dependent variable, here is how we coded its values:

$$Y = \begin{cases} 1, \text{ if the respondent is an internet user} \\ 0, \text{ if the respondent is not an internet user} \end{cases}$$ [8]

The respondents who answered no to question PU_Q01 and those who were marked as valid skip for answering no to the previous question[9] indicating that they have never used the internet are together coded 0. In other words, those who did not use the internet during the past 12 months and those who have never used it were all coded 0 in our data preparation to indicate that they are not internet users.

---

[8] Respondents who did not use the internet during the past 12 months and those who have never used it were all coded 0 in our data preparation to indicate that they are not internet users.

[9] Question EV_Q01 asks the user: Have you ever used the internet from any location for personal non-business use?

The independent variables

Income (inc): The income variable indicates the annual household income of the respondent. The information is recorded in five categories to preserve the respondent's confidentiality. Transforming it to a binary variable by reducing the number of categories will result in loss of information. We prefer to transform it to a continuous variable to get the most information from the data we have. Also, we would like to be consistent with the original model where income was a continuous variable. Since we do not have access to the actual continuous values, we used the five categories to create a continuous variable by assigning the mid-point of each category to the respondents belonging to that category. In doing so, we lose a little variance but all errors are compensated for assuming that income is linearly distributed within each category. Because the fifth category is open ended, we cannot assign a midpoint for it. To represent the 95K and above category, we multiplied 95,000 by 1.5 and assigned the result as the midpoint for that category. After determining the mid-points for each category, we divided them by 1000 because the unit of measurement in the original model was per 1000$. Below is each category with its corresponding continuous value:

1. <= 24,000$ -> 12
2. 24,001 – 39,999 -> 32
3. 40,000 – 59,999 -> 50
4. 60,000 - 94,999 -> 77.5
5. 95,000 + -> 142.5

Age (age): The age of the respondent is also categorized into six categories. Just as we did with the income variable, we used the age categories to create a continuous variable. The sixth category, 65

years and older, is open ended To assign a mid point for this category, we used the average life expectancy in Canada, 81 years [29], as an end value for this category and calculated the mid point accordingly. Below is each category with its corresponding mid-point value:

1. 16 to 24 -> 20.5
2. 25 to 34 -> 30
3. 35 to 44 -> 40
4. 45 to 54 -> 50
5. 55 to 64 -> 60
6. 65+ -> 73

Education (edc): The respondent's highest level of education is categorized into three categories:

1. High-school or less
2. College or some post-secondary education
3. University certificate or degree

In the original model, education was a binary variable where post secondary education was the breakpoint for the two categories. Therefore, we combined the second and third categories together in order to have a binary variable.

$$edc = \begin{cases} 1, \text{ if the respondent has some post - secondary education or higher} \\ 0, \text{ if the respondent has no post - secondary education} \end{cases}$$

Gender (gnd): A binary variable indicating whether the respondent is male or female.

$$gnd = \begin{cases} 1, \text{ if the respondent is a male} \\ 0, \text{ if the respondent is a female} \end{cases}$$

Location (loc): A binary variable indicating whether the respondent resides in and urban or rural area. The location variable in the CIUS dataset is categorized into six categories:

1. Montreal
2. Toronto
3. Vancouver
4. Other Urban
5. Rural Excluding Prince Edward Island
6. Prince Edward Island

We combined categories 1, 2, 3 & 4 to indicate that the respondent resides in an urban area. Category 5 indicates that the respondent resides in a rural area. Category 6 was deleted, as mentioned in the Deleted Cases section.

$$\text{loc} = \begin{cases} 1 \text{ if the respondent resides in an urban area} \\ 0 \text{ if the respondent resides in a rural area} \end{cases}$$

Privacy Concerns (pcs): In the CIUS, this variable was categorized to three categories:

1. Not concerned
2. Concerned
3. Very Concerned

We entered this variable into the models as a continuous variable where an increase in the value indicates an increase in concerns over privacy.

Presence of Children (poc): We transformed the Type of Household variable to a binary variable indicating the presence of unmarried children less than 16 years in the respondent's household. The

variable was categorized to four categories:

4. Single family household with unmarried children under 16
5. Single family household without unmarried children under 16
6. One person households
7. Multifamily Households

Category one was used to present the presence of children, and categories 2 & 3 were combined to present the households without unmarried children under 16 years. Category 4 was deleted as mentioned in the Deleted Cases section.

$$poc = \begin{cases} 1, \text{ if there is a presence of children less than 18 years of age} \\ 0, \text{ if there is no presence of children in the household} \end{cases}$$

4.4 Empirical Models

The models are logistic regression lines that take the socio-economic variables as explanatory variables and quantify their influence on internet use (the dependent variable) by outputting the odds ratio of each variable.

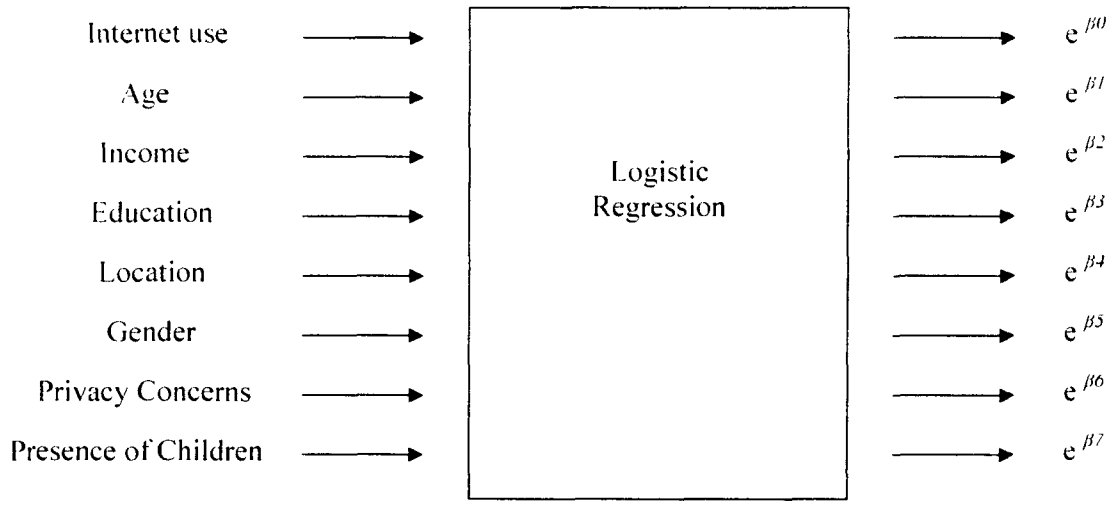


Figure 5: Logistic Regression Input and Output

Figure 5 illustrates the inputs and outputs of the model. "$e^{\beta x}$" is the odds ratio for each of the variables.

Model Specifications:

The standard model for logistic regression

$$Y = \frac{1}{1 + e^{-z}}$$

Where Y is Internet use, e is the base of the natural logarithms, and z is the linear combination which takes different variables for each case:

1- Model for 2005 (Validation model): The model was constructed to compare our results with Noce and McKeown (2007) in order to validate the accuracy of our implementation.

$$z = \beta_0 + \beta_1 \text{age} + \beta_2 \text{inc} + \beta_3 \text{edc} + \beta_4 \text{loc} + \beta_5 \text{gnd} + \beta_6 \text{poc}$$

2- Updated Model for 2007:

$$z = \beta_0 + \beta_1 \text{age} + \beta_2 \text{inc} + \beta_3 \text{edc} + \beta_4 \text{loc} + \beta_5 \text{gnd}$$

3- Model for the Six Regions:

$$z = \beta_0 + \beta_1 \text{age} + \beta_2 \text{inc} + \beta_3 \text{edc} + \beta_4 \text{loc} + \beta_5 \text{gnd} + \beta_6 \text{pcs}$$

where β_0 is the constant and $\beta_{1\ldots6}$ are the coefficients for the respective variables.

4.5 Validation

To validate our implementation, we built a model based on the CIUS 2005 in order to compare it with the original model and assess the accuracy of our model. We applied the same data preparation and model building process to the CIUS 2005 and compared the results we obtained with the results of the original model. We expect to see differences between our model and the original model because the original model included an additional variable, language, and because the original model used the actual values for age and income. However, what we would like to know is how big the difference is in order to asses the accuracy of our model. Table 13 shows a comparison between the original results and our results.

Table 13: Comparison between odds ratios of the Original model and our 2005 model

Variables	Original Results	Our Results	Percentage of Change
Age	0.939	0.941	0.2%
Income	1.01	1.016	0.6%
Education	2.96	2.825	4.6%
Location	1.44	1.42	1.4%
Gender	0.801	0.790	1.4%
Presence of Children	1.05	1.087	3.5%
Constant	7.47	7.15	4.3%
R^2	0.422	0.419	0.7%

The odds ratios for the age and income are very close. This shows that our process of transforming the age and income variables from categorical to continuous variables is highly reliable. There are larger differences in the gender and education variables, but the differences remain quite small. All the variables in the original model were indicated to be significant except Presence of Children. In our model, all the variables were significant on the 5%, but Presence of Children had the least significant among the variables studied[10].

In regression analysis, the R-square value is used to assess the model's fit to the data, which is an assessment of the reliability of the results. For Logistic Regression, the Nagelkerke R^2 value indicates the percentage of the variation in the outcome variable that is explained by the model. The value for a Nagelkerke R^2 is typically much smaller than the value of the R-square for a linear regression model [30]. The Nagelkerke R^2 values of the two models are very close, it is higher for the original model which indicates a better fit. However, this is expected since the original model used the actual continuous values and an additional explanatory variable.

[10] The significance level for the family type is 0.022. See appendix for the complete results of the analysis.

Chapter Five: Analyzing the Results

Since the sample size of every category in this study is at least 2473 observations and the R^2 coefficient is relatively large, we will assume in our analysis that the regression coefficients are significant as confirmed by the SPSS package. Therefore, our analysis in the remainder of the paper will focus on the values of the regression coefficients rather than their significance.

5.1 Comparing Results from 2005 with 2007

After building our models and assessing their reliability, we compare our 2007 model with the original model. Results from 2005 and 2007 are not directly comparable since the 2007 survey included ages 16 and 17 whereas the 2005 survey didn't [25]. Table 14 shows a comparison of the odds ratios of the variables for each model.

Table 14: Comparison between 2005 and 2007

Year	Age	Income	Education	Location	Gender
2005	0.939	1.01	2.96	1.44	0.801
2007	0.935	1.015	3.06	1.31	0.891

For all our results, the observed significance level is less than 0.05. According to Table 14, the influence of the variables is almost stable except for location and gender which have the most notable change. The influence of being in a rural area has decreased by 13%. This could be a result of the efforts to increase the availability of the internet and broadband in rural areas during the two years. The influence of gender has decreased but is still in the same direction. Females still have

higher odds of using the internet than males in Canada, but the odds are now smaller. Consistent with the 2005 results, the presence of children has no significant influence on internet uptake in Canada.

The influence of age is almost the same although the 2007 sample included 16 and 17 year olds which have accounted for a 1% of the 5% increase in internet penetration between 2005 and 2007 [25]. The unchanged odds ratio for age could be explained by an increase in internet uptake among older individuals which would balance the increase among all ages and cause the odds ratio to be stable.

There is also an increase in the influence of education. According to Sciadas [5], the influence of income decreases over time, however our results do not indicate a change in this variable's influence. This could indicate that the influence of income has decreased over time and is now stable as the internet is becoming mature in Canada. It could also mean that more than two years are needed to see a difference in the influence of this variable.

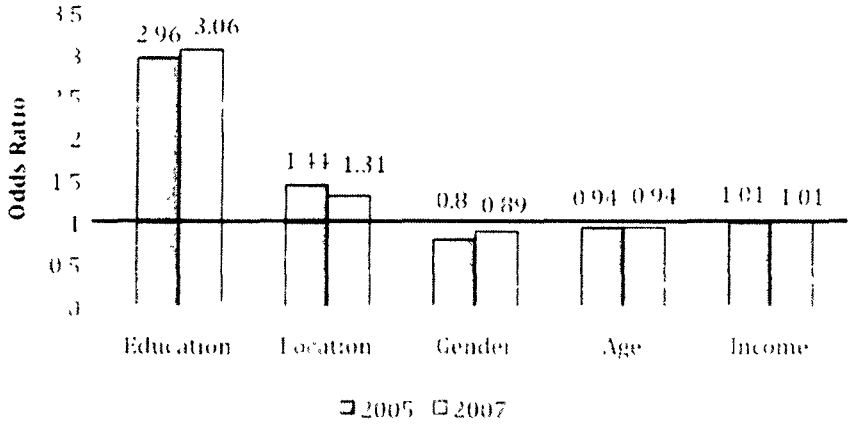


Figure 6: Comparison between 2005 and 2007 Results

Figure 6 shows a comparison between the 2005 and 2007 results. According to Statistics Canada, there was a 5% increase in the internet penetration level between 2005 and 2007. We already know

that a 1% increase came from the respondents aged 16 and 17 [25]. Therefore, according to our model, the remaining 4% came from the increased availability of infrastructure in rural areas.

5.2 Regional Comparison

Table 15: The Odds Ratios for each of the variables per region

Region of Respondent	Age	Income	Education	Location	Gender	Privacy Concern
Atlantic Provinces	0.93	1.02	3.62	1.70	0.723	0.603
Quebec	0.93	1.02	2.42	1.72	0.837*	0.745
Ontario	0.94	1.01	3.56	NS	NS	0.808
Manitoba/ Saskatchewan	0.93	1.01	3.42	1.21*	NS	NS
Alberta	0.94	1.01	3.10	1.70	0.546	0.706
British Colombia	0.94	1.01	2.07	NS	NS	0.814

NS: Not significant

* indicates a significance level larger than 0.01

Table 15 shows the odds ratio for each variable per region. Age, income and education are significant factors that influence internet use in all the six regions. Location, gender and privacy concerns have a significant influence on some of the regions. For every additional year of age, the odds of using the internet decrease by 6% for Ontario, Alberta and BC, and by 7% for the rest of the

regions. For every 1000$ increase in the annual income, the odds of using the internet increase by 2% for the Atlantic Provinces and Quebec, and by 1% for the rest of the regions. This is not surprising since the Atlantic Provinces and Quebec also have the lowest income levels (see table 8).

Location is significant for the Atlantic Provinces, Quebec and Alberta. It also has some significance in Manitoba and Saskatchewan although weaker. The Atlantic Provinces have a larger percentage of rural population and lower broadband availability, which would explain the influence of location. Alberta, however, has a small percentage of rural population and a high broadband availability and it is influenced by the location variable. The Manitoba and Saskatchewan region has a relatively high rural population, yet the location variable does not have a strong influence on this region. Consistent with results from the CIP, our study suggests that Quebec and the Manitoba and Saskatchewan region are both challenged with a rural population, nevertheless, rural residents of Manitoba and Saskatchewan are more likely to be internet users than rural residents of Quebec. This suggests that programs to increase internet penetration in rural areas have been most successful in the Manitoba & Saskatchewan region. Further research is needed to confirm this observation.

Education influences internet use such that the odds of using the internet are 2 times to 3.6 times higher for someone with post secondary education. Quebec and BC have a noticeable lower odds ratio for education when compared to the other regions, but they also have higher educational attainment when compared to the other regions (see table 9). This indicates that individuals with high school or less are more likely to be internet users in Quebec and BC than in other regions.

Although the privacy concerns reported are almost equal among the six regions, their influence on internet use varies. Privacy concerns reduce the odds of using the internet by almost 40% in the Atlantic Provinces. The rest of the regions are affected by 20 to 30% except for Manitoba and Saskatchewan. According to the results, concerns over privacy do not have a significant influence

on internet use in Manitoba and Saskatchewan. This suggests that individuals in this region are experienced internet users.

The odds of using the internet for males decrease by 28% in the Atlantic Provinces, and by 17% in Quebec. An interesting finding is that for males in Alberta the odds of using the internet decrease by almost 50%. For the Atlantic Provinces and Quebec, the numbers are reasonable especially since the result for the 2005 model was quite similar, but for Alberta, the value stands out when compared to the rest of the results. Looking at table 15, the composition of male versus female population in Alberta is consistent with the other regions. Therefore, this is not a composition effect. Further investigation needs to be done to find out the reason behind this result. Figure 7 shows a graphical presentation of the influence of each variable per region.

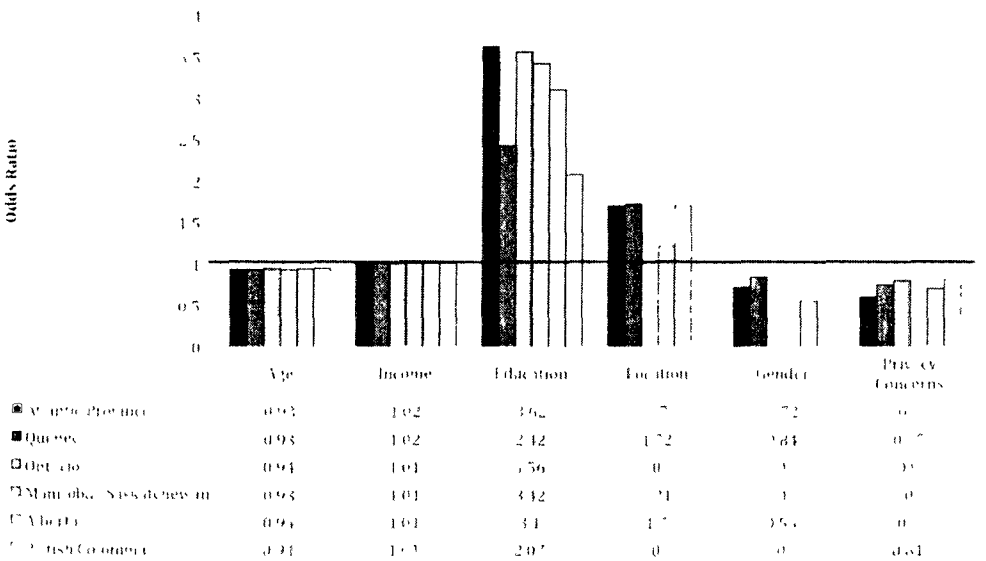


Figure 7: Comparison between the Regions

Looking at the three provinces with the highest penetration rates, we find that the variables have varying influences on each of them. Alberta is strongly affected by the variables, compared to the other regions. This suggests that Alberta has high penetration levels due to the demographics in the province; the higher incomes, the younger population and the lower rural population.

Chapter Six: Discussions and Conclusions

In this research, we studied the socio-economic factors that influence internet penetration in Canada. We built an updated model for the latest data available (for 2007) and compared our results with those of 2005. We used the CIUS 2007 data as our data source. We applied logistic regression to build the models and conduct the analysis. We also performed a regional analysis by building a model for each region and comparing the influence of the variables on each region.

We found that the effect of gender and location (urban vs. rural) on internet use have decreased between 2005 and 2007. However, there is almost no change in the rest of the variables. We already know that there was a 5% increase in internet penetration in Canada between 2005 and 2007, and that 1% of the 5% comes from the inclusion of 16 and 17 year olds in the 2007 survey [25]. Our results suggest that the remaining 4% is a result of the increased availability of infrastructure in rural areas. This also suggests that the initiatives that took place in these two years have been productive and that the digital divide between urban and rural areas is narrowing.

The second objective of this study was to investigate the digital divide among the regions of Canada. The results of our analysis allowed us to interpret the differences among the regions and draw some conclusions. Consistent with previous research, our results suggest that we still need to address supply and demand with a focus on Quebec and the Atlantic Provinces. However, there should be a focus on the Atlantic Provinces when addressing the demand since the education and privacy concern variables had a greater influence in this region (see figure 7).

We know that Quebec and the Atlantic Provinces have lower penetration rates than the rest of the regions and we have found that the variables have a large influence in these two regions. What we did not expect is that Alberta is also influenced strongly by the variables although it has a high

penetration level This suggests that Alberta has high penetration levels due to the demographics in the province, the higher incomes, the younger population and the lower rural population This also suggests that although the penetration rates are high in this province, there is a digital divide in Alberta stronger than that in Ontario and BC, which are the provinces with the highest penetration rates The divide between urban and rural residents is as strong as that in the Atlantic Provinces and Quebec

The Manitoba and Saskatchewan region has a large rural population and broadband penetration is on the lower side when compared to the other regions Yet, it is not strongly affected by the location variable, especially when considering the large rural population in this region (see table 5) This suggests that the initiatives that took place to increase internet penetration in this region have been the most successful. Also, this is the only region that is not affected by privacy concerns, which indicates that individuals in this region are experienced internet users

We have encountered some difficulties in conducting this research which limited the results we could achieve The language factor, or proficiency in English, is an important factor for internet use in Canada, but we could not add it to our study due to lack of data access We were also interested in studying the effect of relative content on internet use in Canada but the design of the survey restricted us from including this as a variable in the model

6.1 Recommendations

Based on the study conducted in this thesis and our observations on the influence of the variables studied, we make the following recommendations

- The divide between urban and rural areas is narrowing, which indicates that the initiatives that took place have been productive and should continue for further progress

- The provinces least influenced by the demographic variables are not necessarily the ones with the highest penetration rates. Therefore, a mere look at the penetration rates across the regions is not enough to assess the connectivity of the region and the digital divide within.
- There should be a focus on the Atlantic Provinces when addressing the demand for internet use.
- Although Alberta is one of the regions with high penetration levels, our results suggest that there is a divide among the residents of this province based on the variables studied in this research. We believe that the penetration rate in this province masks the disparities among the citizens. Future initiatives should have a focus on this province along with the Atlantic Provinces and Quebec.
- Although we cannot confirm it, the results suggest that the Manitoba and Saskatchewan region is the most successful to overcome the influence of rurality on internet use. Future efforts to increase internet uptake in rural areas should consider this observation and draw success lessons from the programs of this region.

6.2 Future Work

- The results of the original model (2005) and our model (2007) both indicate that the influence of education is approximately twice as large as that of location. This indicates that the demand barrier is more significant than the supply barrier. Consequently, we believe that future efforts in this area should focus on the dual digital divide (i.e. the divide between those who are interested in using the internet and those who are not)

- The CIUS 2009 has already taken place and results will be released during 2010. Comparing the 2009 results with 2005 will provide more insight into the pattern of the variables' influence on internet penetration in Canada

References

[1] Selwyn, N. (2004). Reconsidering Political and Popular understandings of the Digital Divide. *New Media and Society*, 6(3), 341-362.

[2] Frieden, R. (2005). Lessons from broadband development in Canada, Japan, Korea and the United States. *Telecommunications Policy*, 29(8), 595-613.

[3] Middleton, C. A., Sorensen C. (2005). How Connected are Canadians? Inequities in Canadian Households' Internet Access. *Canadian Journal of Communication*, 30, 463-483.

[4] Chinn, M. D., & Fairlie, R. W. (2007). The determinants of the global digital divide: A cross-country analysis of computer and internet penetration. *Oxford Economic Papers*, 59(1), 16-44.

[5] Sciadas, George (2002). Unveiling the digital divide.Connectedness Series [Statistics Canada, Cat. No. 56F0004MIE, No. 7.].

[6] International Telecommunication Union ITU (2007). "Measuring the Information Society: ICT Opportunity Index and World Telecommunication/ICT Indicators". Geneva.

[7] Statistics Canada, http://www.statcan.gc.ca/pub/56f0004m/2005012/ch2-eng.htm accessed on September 2nd, 2009.

[8] Noce, A., & McKeown, L. (2007). A new benchmark for internet use: A logistic modeling of factors influencing internet use in Canada, 2005. *Government Information Quarterly*, 25(3), 462-476.

[9] McKeown, L., Noce, A., & Czerny, P. (2007). Factors Associated with Internet Use: Does Rurality Matter? Rural and Small Town Canada Analysis Bulletin [Statistics Canada, Cat. No. 21-006-XIE, Vol. 7, No. 3]

[10] Belanger, F., & Carter, L. (2009). The Impact of the Digital Divide on e-Government Use. *Communications of the ACM*, 52(4), 132-135.

[11] Cullen, R. (2001). Addressing the digital divide. *Online Information Review*, 25, 311-320.

[12] Zamaria, C., & Fletcher, F. (2005). Canada online! The Internet, Media and Emerging Technologies: Uses, Attitudes, Trends and International Comparisons 2007. Toronto, ON: Canadian Internet Project.

[13] International Telecommunication Union ITU (2003). "Promoting Broadband: the case of Canada" Workshop on Promoting Broadband. Geneva.

[14] Industry Canada, http://www.ic.gc.ca/eic/site/cap-pac.nsf/eng/00006.html accessed on December 2nd, 2009.

[15] Rideout, V., (2000). Public Access to the Internet and the Canadian Digital Divide. *Canadian Journal of Information and Library Science*, 25.

[16] http://www.servicealberta.gov.ab.ca/AlbertaSuperNet.cfm accessed November, 2009.

[17] http://www.network.gov.bc.ca/about.htm accessed November, 2009.

[18] http://www.ito.gov.sk.ca/communitynet/ accessed August, 2009.

[19] http://www.infrastructure.mb.ca/en/cmip.html

[20] http://www.gnb.ca/cnb/news/snb/2005e0052sn.htm accessed November, 2009.

[21] http://www.gnb.ca/cnb/news/bnb/2009e0102bn.htm accessed August, 2009.

[22] http://www.meq.gouv.qc.ca/lancement/villagesbranches/

[23] Quebec On-Line: Promoting the use of the Internet and the Development of e-Commerce

[24] http://www.rrq.gouv.qc.ca/en/Pages/accueil.aspx

[25] Statistics Canada: The Daily, http://www.statcan.gc.ca/daily-quotidien/080612/dq080612b-eng.htm accessed March, 2009.

[26] Statistics Canada: CIUS, http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=4432&lang=en&db=imdb&adm=8&dis=2

[27] Odesi: http://search1.odesi.ca/ Accessed November 2009.

[28] Microdata User Guide, http://prod.library.utoronto.ca:8090/datalib/codebooks/cstdli/hius/2007/cius2007gid.pdf Accessed November 2009.

[29] CIA World Fact Book, https://www.cia.gov/library/publications/the-world-factbook/rankorder/2102rank.html Accessed November 2009.

[30] Norusis, 2005. *SPSS 13.0 Statistical Procedures Companion.* Prentice Hall, Inc.

Appendix

SPSS Tables:

1) CIUS 2005

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GCAGEGR6	-.070	.001	5052.455	1	.000	.933
	Constant	4.162	.053	6255.083	1	.000	64.184
Step 2[b]	GCAGEGR6	-.063	.001	4036.556	1	.000	.939
	G_HQUINT	.019	.000	2038.596	1	.000	1.019
	Constant	2.705	.057	2216.831	1	.000	14.952
Step 3[c]	GCAGEGR6	-.061	.001	3662.281	1	.000	.941
	G_CEDUC	1.051	.031	1131.078	1	.000	2.860
	G_HQUINT	.016	.000	1450.798	1	.000	1.016
	Constant	2.165	.060	1322.986	1	.000	8.712
Step 4[d]	G_URBRUR	.354	.035	103.215	1	.000	1.425
	GCAGEGR6	-.061	.001	3640.422	1	.000	.941
	G_CEDUC	1.034	.031	1088.531	1	.000	2.812
	G_HQUINT	.016	.000	1393.701	1	.000	1.016
	Constant	1.928	.064	914.656	1	.000	6.878
Step 5[e]	G_URBRUR	.351	.035	101.131	1	.000	1.420
	GCAGEGR6	-.061	.001	3659.153	1	.000	.940
	CSEX	-.239	.031	58.342	1	.000	.788
	G_CEDUC	1.037	.031	1091.291	1	.000	2.820
	G_HQUINT	.016	.000	1424.833	1	.000	1.016
	Constant	2.043	.066	966.440	1	.000	7.714
Step 6[f]	G_URBRUR	.352	.035	101.735	1	.000	1.422
	GCAGEGR6	-.060	.001	2992.063	1	.000	.941
	CSEX	-.235	.031	56.653	1	.000	.790
	G_CEDUC	1.039	.031	1093.901	1	.000	2.825
	GFAMTYPE	.084	.036	5.267	1	.022	1.087
	G_HQUINT	.016	.000	1390.185	1	.000	1.016
	Constant	1.967	.073	719.340	1	.000	7.150

a. Variable(s) entered on step 1: GCAGEGR6.
b. Variable(s) entered on step 2: G_HQUINT.
c. Variable(s) entered on step 3: G_CEDUC.
d. Variable(s) entered on step 4: G_URBRUR.
e. Variable(s) entered on step 5: CSEX.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GCAGEGR6	-.070	.001	5052.455	1	.000	.933
	Constant	4.162	.053	6255.083	1	.000	64.184
Step 2[b]	GCAGEGR6	-.063	.001	4036.556	1	.000	.939
	G_HQUINT	.019	.000	2038.596	1	.000	1.019
	Constant	2.705	.057	2216.831	1	.000	14.952
Step 3[c]	GCAGEGR6	-.061	.001	3662.281	1	.000	.941
	G_CEDUC	1.051	.031	1131.078	1	.000	2.860
	G_HQUINT	.016	.000	1450.798	1	.000	1.016
	Constant	2.165	.060	1322.986	1	.000	8.712
Step 4[d]	G_URBRUR	.354	.035	103.215	1	.000	1.425
	GCAGEGR6	-.061	.001	3640.422	1	.000	.941
	G_CEDUC	1.034	.031	1088.531	1	.000	2.812
	G_HQUINT	.016	.000	1393.701	1	.000	1.016
	Constant	1.928	.064	914.656	1	.000	6.878
Step 5[e]	G_URBRUR	.351	.035	101.131	1	.000	1.420
	GCAGEGR6	-.061	.001	3659.153	1	.000	.940
	CSEX	-.239	.031	58.342	1	.000	.788
	G_CEDUC	1.037	.031	1091.291	1	.000	2.820
	G_HQUINT	.016	.000	1424.833	1	.000	1.016
	Constant	2.043	.066	966.440	1	.000	7.714
Step 6[f]	G_URBRUR	.352	.035	101.735	1	.000	1.422
	GCAGEGR6	-.060	.001	2992.063	1	.000	.941
	CSEX	-.235	.031	56.653	1	.000	.790
	G_CEDUC	1.039	.031	1093.901	1	.000	2.825
	GFAMTYPE	.084	.036	5.267	1	.022	1.087
	G_HQUINT	.016	.000	1390.185	1	.000	1.016
	Constant	1.967	.073	719.340	1	.000	7.150

a. Variable(s) entered on step 1: GCAGEGR6.
b. Variable(s) entered on step 2: G_HQUINT.
c. Variable(s) entered on step 3: G_CEDUC.
d. Variable(s) entered on step 4: G_URBRUR.
e. Variable(s) entered on step 5: CSEX.
f. Variable(s) entered on step 6: GFAMTYPE.

2) CIUS 2007

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GCAGEGR6	-.075	.001	4651.922	1	.000	.928
	Constant	4.744	.061	6072.353	1	.000	114.943
Step 2[b]	GCAGEGR6	-.068	.001	3783.936	1	.000	.935
	G_HQUINT	.017	.000	1493.657	1	.000	1.017
	Constant	3.322	.066	2564.662	1	.000	27.707
Step 3[c]	GCAGEGR6	-.067	.001	3646.845	1	.000	.935
	G_CEDUC	1.125	.035	1024.973	1	.000	3.081
	G_HQUINT	.015	.000	1073.584	1	.000	1.015
	Constant	2.846	.067	1820.047	1	.000	17.214
Step 4[d]	G_URBRUR	.274	.040	48.076	1	.000	1.315
	GCAGEGR6	-.067	.001	3627.341	1	.000	.935
	G_CEDUC	1.116	.035	1004.647	1	.000	3.052
	G_HQUINT	.014	.000	1037.672	1	.000	1.015
	Constant	2.658	.072	1370.616	1	.000	14.262
Step 5[e]	G_URBRUR	.273	.040	47.740	1	.000	1.314
	GCAGEGR6	-.068	.001	3632.397	1	.000	.935
	CSEX	-.115	.035	10.895	1	.001	.891
	G_CEDUC	1.118	.035	1007.892	1	.000	3.060
	G_HQUINT	.015	.000	1045.219	1	.000	1.015
	Constant	2.714	.074	1347.084	1	.000	15.095

a. Variable(s) entered on step 1: GCAGEGR6.
b. Variable(s) entered on step 2: G_HQUINT.
c. Variable(s) entered on step 3: G_CEDUC.
d. Variable(s) entered on step 4: G_URBRUR.
e. Variable(s) entered on step 5: CSEX.

3) Atlantic Provinces

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GCAGEGR6	-.079	.004	469.729	1	.000	.924
	Constant	4.675	.198	559.541	1	.000	107.257
Step 2[b]	GCAGEGR6	-.079	.004	443.783	1	.000	.924
	G_CEDUC	1.570	.111	201.382	1	.000	4.808
	Constant	3.844	.199	371.303	1	.000	46.699
Step 3[c]	GCAGEGR6	-.071	.004	347.647	1	.000	.931
	G_CEDUC	1.296	.115	126.151	1	.000	3.653
	G_HQUINT	.018	.002	105.793	1	.000	1.018
	Constant	2.645	.220	144.180	1	.000	14.087
Step 4[d]	GCAGEGR6	-.068	.004	313.740	1	.000	.934
	G_CEDUC	1.309	.117	125.644	1	.000	3.701
	G_HQUINT	.018	.002	104.554	1	.000	1.018
	PS_Q01	-.482	.074	42.715	1	.000	.618
	Constant	3.570	.271	174.028	1	.000	35.515
Step 5[e]	G_URBRUR	.529	.116	20.903	1	.000	1.698
	GCAGEGR6	-.069	.004	310.188	1	.000	.934
	G_CEDUC	1.291	.117	120.891	1	.000	3.635
	G_HQUINT	.017	.002	95.777	1	.000	1.017
	PS_Q01	-.488	.074	43.333	1	.000	.614
	Constant	3.367	.275	149.994	1	.000	28.987
Step 6[f]	G_URBRUR	.529	.116	20.795	1	.000	1.698
	GCAGEGR6	-.069	.004	309.102	1	.000	.933
	CSEX	-.324	.117	7.693	1	.006	.723
	G_CEDUC	1.286	.118	119.597	1	.000	3.620
	G_HQUINT	.018	.002	98.633	1	.000	1.018
	PS_Q01	-.506	.075	45.971	1	.000	.603
	Constant	3.551	.285	155.166	1	.000	34.835

a. Variable(s) entered on step 1: GCAGEGR6.
b. Variable(s) entered on step 2: G_CEDUC.
c. Variable(s) entered on step 3: G_HQUINT.
d. Variable(s) entered on step 4: PS_Q01.
e. Variable(s) entered on step 5: G_URBRUR.
f. Variable(s) entered on step 6: CSEX.

4) Quebec

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GCAGEGR6	-.082	.003	940.013	1	.000	.922
	Constant	5.036	.147	1176.607	1	.000	153.895
Step 2[b]	GCAGEGR6	-.075	.003	775.500	1	.000	.928
	G_HQUINT	.021	.001	299.191	1	.000	1.021
	Constant	3.577	.156	523.392	1	.000	35.761
Step 3[c]	GCAGEGR6	-.072	.003	696.960	1	.000	.931
	G_CEDUC	.865	.082	111.573	1	.000	2.375
	G_HQUINT	.018	.001	221.658	1	.000	1.019
	Constant	3.061	.162	356.655	1	.000	21.340
Step 4[d]	G_URBRUR	.536	.088	36.752	1	.000	1.709
	GCAGEGR6	-.072	.003	691.266	1	.000	.931
	G_CEDUC	.855	.082	108.055	1	.000	2.352
	G_HQUINT	.018	.001	206.398	1	.000	1.018
	Constant	2.719	.171	252.627	1	.000	15.172
Step 5[e]	G_URBRUR	.542	.089	37.440	1	.000	1.720
	GCAGEGR6	-.072	.003	677.821	1	.000	.931
	G_CEDUC	.878	.083	112.432	1	.000	2.407
	G_HQUINT	.018	.001	209.852	1	.000	1.018
	PS_Q01	-.286	.052	30.393	1	.000	.751
	Constant	3.308	.206	258.468	1	.000	27.336
Step 6[f]	G_URBRUR	.540	.089	37.071	1	.000	1.716
	GCAGEGR6	-.072	.003	679.385	1	.000	.930
	CSEX	-.178	.082	4.683	1	.030	.837
	G_CEDUC	.884	.083	113.671	1	.000	2.421
	G_HQUINT	.018	.001	213.009	1	.000	1.019
	PS_Q01	-.295	.052	31.948	1	.000	.745
	Constant	3.415	.212	258.442	1	.000	30.428

a. Variable(s) entered on step 1: GCAGEGR6.
b. Variable(s) entered on step 2: G_HQUINT.
c. Variable(s) entered on step 3: G_CEDUC.
d. Variable(s) entered on step 4: G_URBRUR.
e. Variable(s) entered on step 5: PS_Q01.
f. Variable(s) entered on step 6: CSEX.

5) Ontario

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GCAGEGR6	-.071	.002	1129.666	1	.000	.931
	Constant	4.730	.118	1618.685	1	.000	113.300
Step 2[b]	GCAGEGR6	-.070	.002	1070.944	1	.000	.932
	G_CEDUC	1.457	.067	479.382	1	.000	4.295
	Constant	3.913	.119	1087.655	1	.000	50.067
Step 3[c]	GCAGEGR6	-.065	.002	909.762	1	.000	.937
	G_CEDUC	1.258	.069	331.582	1	.000	3.517
	G_HQUINT	.014	.001	264.277	1	.000	1.014
	Constant	2.819	.129	479.263	1	.000	16.765
Step 4[d]	GCAGEGR6	-.063	.002	856.060	1	.000	.938
	G_CEDUC	1.269	.069	335.867	1	.000	3.558
	G_HQUINT	.014	.001	265.404	1	.000	1.014
	PS_Q01	-.213	.043	24.192	1	.000	.808
	Constant	3.217	.154	435.458	1	.000	24.947

a. Variable(s) entered on step 1: GCAGEGR6.
b. Variable(s) entered on step 2: G_CEDUC.
c. Variable(s) entered on step 3: G_HQUINT.
d. Variable(s) entered on step 4: PS_Q01.

6) Manitoba/Saskatchewan

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GCAGEGR6	-.076	.003	675.111	1	.000	.927
	Constant	4.778	.162	868.896	1	.000	118.817
Step 2[b]	GCAGEGR6	-.076	.003	668.369	1	.000	.926
	G_CEDUC	1.396	.092	229.429	1	.000	4.040
	Constant	4.120	.161	655.586	1	.000	61.545
Step 3[c]	GCAGEGR6	-.072	.003	578.805	1	.000	.931
	G_CEDUC	1.252	.095	172.603	1	.000	3.498
	G_HQUINT	.014	.001	131.487	1	.000	1.014
	Constant	3.117	.174	321.163	1	.000	22.573
Step 4[d]	G_URBRUR	.190	.095	3.979	1	.046	1.209
	GCAGEGR6	-.071	.003	567.142	1	.000	.931
	G_CEDUC	1.229	.096	164.121	1	.000	3.419
	G_HQUINT	.014	.001	128.975	1	.000	1.014
	Constant	2.997	.183	266.824	1	.000	20.031

a. Variable(s) entered on step 1: GCAGEGR6.
b. Variable(s) entered on step 2: G_CEDUC.
c. Variable(s) entered on step 3: G_HQUINT.
d. Variable(s) entered on step 4: G_URBRUR.

7) Alberta

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GCAGEGR6	-.070	.004	362.733	1	.000	.932
	Constant	4.710	.200	554.216	1	.000	111.016
Step 2[b]	GCAGEGR6	-.072	.004	367.637	1	.000	.930
	G_CEDUC	1.265	.117	117.502	1	.000	3.542
	Constant	4.155	.201	427.017	1	.000	63.767
Step 3[c]	GCAGEGR6	-.065	.004	297.506	1	.000	.937
	G_CEDUC	1.112	.120	85.609	1	.000	3.041
	G_HQUINT	.012	.001	77.343	1	.000	1.012
	Constant	3.009	.225	178.450	1	.000	20.260
Step 4[d]	GCAGEGR6	-.066	.004	301.146	1	.000	.936
	CSEX	-.556	.122	20.648	1	.000	.574
	G_CEDUC	1.124	.121	86.119	1	.000	3.078
	G_HQUINT	.013	.001	87.020	1	.000	1.014
	Constant	3.285	.236	193.587	1	.000	26.722
Step 5[e]	GCAGEGR6	-.065	.004	286.823	1	.000	.937
	CSEX	-.601	.124	23.660	1	.000	.548
	G_CEDUC	1.163	.122	90.377	1	.000	3.200
	G_HQUINT	.014	.001	90.480	1	.000	1.014
	PS_Q01	-.358	.078	20.756	1	.000	.699
	Constant	3.997	.290	190.267	1	.000	54.459
Step 6[f]	G_URBRUR	.529	.127	17.346	1	.000	1.698
	GCAGEGR6	-.065	.004	280.761	1	.000	.937
	CSEX	-.606	.124	23.741	1	.000	.546
	G_CEDUC	1.129	.123	84.391	1	.000	3.091
	G_HQUINT	.014	.001	91.687	1	.000	1.014
	PS_Q01	-.348	.079	19.445	1	.000	.706
	Constant	3.620	.303	142.610	1	.000	37.325

a. Variable(s) entered on step 1: GCAGEGR6.
b. Variable(s) entered on step 2: G_CEDUC.
c. Variable(s) entered on step 3: G_HQUINT.
d. Variable(s) entered on step 4: CSEX.
e. Variable(s) entered on step 5: PS_Q01.
f. Variable(s) entered on step 6: G_URBRUR.

8) British Colombia

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GCAGEGR6	-.072	.004	397.187	1	.000	.931
	Constant	5.015	.206	591.552	1	.000	150.646
Step 2[b]	GCAGEGR6	-.065	.004	341.455	1	.000	.937
	G_HQUINT	.014	.001	93.365	1	.000	1.014
	Constant	3.787	.222	290.489	1	.000	44.125
Step 3[c]	GCAGEGR6	-.065	.004	345.471	1	.000	.937
	G_CEDUC	.723	.110	43.090	1	.000	2.060
	G_HQUINT	.013	.001	80.765	1	.000	1.013
	Constant	3.434	.225	233.729	1	.000	31.007
Step 4[d]	GCAGEGR6	-.064	.004	324.689	1	.000	.938
	G_CEDUC	.726	.110	43.416	1	.000	2.068
	G_HQUINT	.013	.001	80.803	1	.000	1.013
	PS_Q01	-.206	.072	8.266	1	.004	.814
	Constant	3.821	.265	207.847	1	.000	45.644

a. Variable(s) entered on step 1: GCAGEGR6.
b. Variable(s) entered on step 2: G_HQUINT.
c. Variable(s) entered on step 3: G_CEDUC.
d. Variable(s) entered on step 4: PS_Q01.