Utilization of General Practitioners' Services in Canada and the United States: A Quantile Regression for Counts Analysis

*Chunping Liu**

Department of Economics

University of Guelph

Abstract

This paper uses data from the Joint Canada/United States Health Survey (JCUSH) to investigate the determinants of General Practitioner (GP) utilization in the two countries. It applies the quantile regression for count data model (QRCM) proposed by Machado and Santos Silva (2005) to an equation derived from Grossman's model of the demand for health capital to investigate factors affecting GP visits across the two countries, and whether changes in the values of the explanatory variables have different effects on GP utilization at different quantiles of the distribution. It discusses differences and similarities in the factors affecting GP utilization in the two countries, and also compares the implications drawn from the quantile regression approach with those from Two-Part Model (TPM) which is commonly used in the health economics literature. The results suggest that QRCM provides more information on how various factors affect utilization of GP services than TPM does, providing information not only on how the distribution of the dependent variable shifts when the value of an explanatory variable changes, but also on whether the shape of the distribution changes at different quantiles conditional on explanatory variables. The QRCM results in this paper show that sex, self-assessed health status, the health utility index, having a regular doctor and number of chronic diseases are key determinants of GP visits, and that they shift the distribution of the number of GP visits more at high quantiles than at low quantiles. Except for having a regular doctor, these factors have bigger impacts for Canadians than for Americans on GP utilization across all the quantiles. The effect on GP utilization of having a regular GP seems to be the same, at all quantiles, across the two countries. Other variables including insurance variables, age, immigrant, smoking dummies and Body Mass Index dummies are also significant at some quantiles of GP utilization.

Keywords: Quantile Regression, Count Data Model, Two-part Model, Physicians' Services Utilization.

---

*Department of Economics, University of Guelph, Guelph, Ontario, Canada N1G 2W1. Tel: (519)824-4120. Fax: (519) 763-8497. E-mail: liuc@uoguelph.ca and

I would like to thank Brian Ferguson, Yiguo Sun, Audrey Laporte, Miana Plesca and participants at the Atlantic Canada Economics Association 2007 conference for their comments and suggestions. All remaining errors are on my own.

I. Introduction

The idea of analyzing the difference in health care utilization between Canada and the U.S. is based on the institutional differences in health care systems in the two countries. The major differences lie in health insurance coverage and the way health service is delivered: the Canadian system provides comprehensive health coverage universally to Canadians, with private health insurance (most notably for outpatient pharmaceuticals, which are not covered under Canadian Medicare) acting as a supplement to public health insurance. In terms of health service delivery, the Canadian system is characterized by strong government intervention with a limited private sector role. In the U.S., the health system relies heavily on the private sector. There are government health insurance programs designed for the poor (Medicaid), and seniors or handicapped persons (Medicare). It is not compulsory, in most states, for Americans to carry health care insurance. They can obtain private health insurance from their employers, or purchase it themselves. Denavas-Walt et al. (2007) reports that 15.8% of the population, i.e. roughly 47.0 million people, in the U.S. lacks health insurance and that the percentage of people covered by employment-based health insurance or government health programs was lower in 2006 than in 2005.

There is an extensive literature comparing the health systems in the two countries, including studies of health care utilization. Health care utilization includes three major aspects: physicians, hospitals and drugs. This paper focuses on the utilization of General Practitioner (GP) services .

The central problem investigated in this paper is characteristics of users' behavior across the distribution of GP consultations. . Since the number of GP services is discrete data, heavily concentrated at zero services, the previous literature has used count data models. The commonly used count data models include the Poisson Model, the negative binomial Model, and extensions such as the Two-part model (TPM) and the Latent Class model (LCM). All of these models impose certain assumptions on the distributions of number of GP visits, i.e., they assume the distribution of the number of GP visits satisfies either Poisson or negative binomial distribution. In addition, they all provide information of the distribution of the number of GP visits only around its mean, but not at other locations of the distribution conditional on explanatory variables.

The Quantile regression for counts model (QRCM), proposed by Machado and Santos Silva (2005) is an innovative count model based on Quantile Regression (QR), which differs from the above count data models used in the previous literature in that it does not impose any assumption on the distribution of the number of GP visits and it has the flexibility of investigating how the distribution of the number of GP visits changes at different conditional quantiles. Traditional regression approaches estimate how the conditional mean changes with the change in explanatory variables, while QR provides a broader view, looking at how the distribution of the dependent variable changes at various conditional quantiles. As long as there is a change in the shape of the distribution of the dependent variable at different quantiles when there is a change on the value of covariates, it is worth applying QR. To many policy makers, the stories at the two tails of the distribution of the dependent variable are often of particular interests in terms of policy implications. Part of the appeal of QR is that it provides us with the information both on the shift of the distribution and on the change of the shape of the distribution of the number of GP visits conditional on explanatory variables. QR estimates the conditional quantile by minimizing an objective function of symmetrically or asymmetrically weighted absolute deviations between the observations and the estimated conditional quantiles, and it therefore requires the dependent variable to be continuous in order to solve the optimization problem. A potential problem with QR on the GP's service utilization comes from the discreteness of the number of GP visits. The QRCM approach proposed by Machado and Santos Silva (2005) tackles this discrete dependent variable problem by artificially introducing an extra continuous noise term to smooth the discontinuous dependent variable and then applying QR to the smoothened dependent variable.

In this paper, I employ QRCM to investigate the characteristics of GP utilization at different conditional quantiles. I address two questions: (i) What are the determinants of GP visits in the two countries? (ii) Is there any difference in the determinants of GP utilization across different quantiles in and across the two countries?. The theoretical model is based on Grossman's (1972a, 1972b) health capital model of the demand for health care, which provides a framework linking human capital with health care utilization.

This paper contributes to health care utilization literature in three aspects. First, it provides an application of QR on counts models in that QR provides information on not only the shifts, but also the changes of shape of the distribution of the dependent variable at different quantiles; while most literature focused only on the model based on users and non-users, and/or only on users. Second, it provides evidence that QRCM is one appealing approach for count data model beyond TPM, a popular count data model, in health care utilization analysis. Third, it provides evidence on the differences in GPs service utilization between Canada and the U.S.

The QRCM results show that sex, self-assessed health status, health utility index, having a regular doctor and number of chronic diseases are key determinants of GP visits. They indicate that changes in the values of the explanatory variables change the shape of the distribution of utilization as well as changing the mean value of that distribution. For many of the variables considered here, the shifts of the distribution are bigger at high quantiles than at low quantiles.

The marginal effects or the shifts of the distribution of the number of GP consultations are bigger in Canada than in the U.S. for sex, self-assessed health status and health utility index, but the same in the two countries for on having a regular doctor. Insurance variables increase GP's service utilization at different quantiles in the two countries.

The remainder of the paper is presented as follows. Section II reviews literature on GP utilization and Joint Canada/United States Survey of Health (JCUSH). In section III, a theoretical Grossman-model-based framework and econometric QRCM are introduced. Section IV elaborates the dataset, model and variables used in this paper. Section V presents the result and discussion. Concluding remarks are given in section VI.

II. Literature Review

I briefly review the physician utilization literature and JCUSH-related research in this section. In particular, I summarize the results on income and insurance because the income effect is often of concern to health policy makers, and health insurance coverage is the most important institutional difference between the health systems in Canada and the U.S.

The majority of the health care utilization literature focuses on physician utilization including GP visits, specialist visits and other health professional visits. In physician utilization research, topics range from cross-country comparison, utilization inequality, the influence of specific determinants including insurance and income, health policy or economic phenomenon effect, to econometric methodology exploration are covered. Results from the former literature show that determinants such as smoking, drinking and Body Mass Index (BMI) variables often exhibit mixed effects on physician service utilization. Demographic and socioeconomic variables often play considerable roles, and variables like number of chronic problems, self-assessed health status, and health utility index consistently show significant effect on physician's service utilization.

Health policy makers often seek answers to such questions as whether the high income population will utilize health care services more than the low income group and whether income is a barrier to impede poor population from using health care utilization. Deb and Trivedi (1997; 2002) do not find family income to affect health care utilization. Dunlop et al. (2000), however, shows that even though income does not directly impede access to health care services, after adjusting for difference in health need, Canadians with lower incomes visit specialists at a lower rate than those with moderate or high incomes. Family income is positively correlated to GP visits in Lahiri and Xing (2004) and Atella et al. (2004). Jimenez-Martin et al. (2004) find income explains a fraction of the variability in the health service demand across EU countries. The effect of family income on health care utilization from former literature is therefore uncertain.

Health insurance is the major institutional difference between Canada and the U.S. It is a heavily investigated variable in health care utilization literature. A large literature has investigated the effect of insurance in different countries and health insurance is often considered to be positively related with the health care utilization. Deb and Trivedi (1997) find that health insurance leads to different physician utilizations for the high- and low-use groups in the U.S. Stabile (2001) also finds that in Canada, individuals who hold private supplemental insurance use more GP service than individuals without such insurance. Rodriguez and Stoyanova (2004) shows insurance access is the main determinant of GP service in Spain.

I use the dataset from the JCUSH in this paper. Previous literature on JCUSH is mostly analysis based on statistical description. Sanmartin et al. (2006) compares health and health care use in Canada and the U.S. and concludes that health status appears to be relatively similar in the two countries and there exists income-related disparities especially in the U.S.. Lasser et al. (2006) uses multivariate analysis to compare health status, access to care, and utilization of medical services in two countries. The conclusion is that Americans are less able to access health care than Canadians due to the institutional difference in health systems in the two countries. Armstrong et al. (2006) investigates the socioeconomic, demographic and health status factors in two countries and finds the correlations between the investigated factors and health are relatively stronger in the U.S. They suggest the different correlations in two countries are the outcome of different health care access.

III. Methodology

III.I Theoretical Model: A Grossman-Model-Based framework

In this paper, I fit a model of health care utilization based on Grossman's human capital model of the demand for health. Grossman (1972a, 1972b) exploits the distinction between health as an output and medical care as an input theoretically and empirically. His model provides an explanation of the demand for medical goods and services. Grossman (1999) has detailed discussion on theoretical extensions of Grossman (1972), theoretical predictions and some empirical research that tests the prediction of his model.

The central idea of Grossman's model is to view health as a durable capital stock that yields an output of healthy time. Individuals inherit an initial amount of health stock that depreciates with age and can be increased by investment. Traditional demand theory assumes that consumers choose a combination of goods and services that maximizes their utility subject to an income or resource constraint, i.e., expenditures on goods and services cannot exceed income. Grossman's model is based on this theory. The framework of Grossman model can be simplified using the following equations.

$$U = U(H_t, X_t) \tag{1}$$

$$H_{t+1} = I(M_t; E) + (1-\delta_t)H_t \tag{2}$$

$$Y_t = p * X_t + q * M_t \tag{3}$$

Where t=0,1,…, n, U is the utility, H_t is the stock of health at age t or in time period t, I_t is gross investment on health, M_t is a vector of inputs (goods, including medical care, on which our discussion focuses) purchased in the market that contribute to gross investment in health, X_t is a vector of goods purchased other than M_t, assumed not to affect health, E is consumer's stock of knowledge or human capital exclusive of health capital, Y_t is the disposable income, p and q are the prices for medical care and consumer goods, which are assumed exogenous and fixed.

In order to maintain a target health state H_{t+1}, a certain amount of medical care M needs to be purchased. The more M the individual consumes, the better the state of health, keeping other situations including age, education, life style etc. the same. We can express equation (2) one period behind as $H_t = I(M_{t-1}; E) + (1-\delta_{t-1})H_{t-1}$ and plug it into (2). By repeated lagging and substitution, we can express H_{t+1} as an accumulation of the depreciated past investments on health.

Equation (1) is individual's budget constraint. Individuals spend their disposable income on purchasing Medical services M and other goods X. The optimal amount of M and X purchased is based on the maximizing individual's life time utility subject to budget constraint. This situation is illustrated in figure 1. We can rewrite equation (3) as $X_t = \frac{Y_t}{p} - \frac{q}{p} M_t$ and plug it into equation (1). The Utility function is thus expressed as:

$$U = U(H_t, \frac{Y_t}{p} - \frac{q}{p} M_t) \tag{4}$$

M_t is therefore associated with p, q, Y_t, H_t and E through the utility function and the budget constraint. The optimal M_t is decided by many factors such as p, q, Y_t, H_t and E. Based on the above analysis, we can express M_t as a function of these factors:

$$M_t = f(p, q, Y_t, H_t; E) \tag{5}$$

Equation (5) is the central problem investigated in this paper. Education E and disposable income Y_t are variables easy to measure. However, health capital H_t is hard to measure. I use number of chronic diseases as one measurement of health capital. Self-assessed health status, the health utility index[1], and BMI are also used as proxies for health capital. Insurance variables affect the prices for medical care (p) and aggregated other goods (q) on the market.

From the dataset used in this paper, around 60% of the respondents (low-users) use two or less GP service and around 10% of the respondents (high-users) use more than six physician consultations. Graphically, most respondents (low GP users) cluster at the low quantiles of M and few respondents (high GP users) are at high quantiles of M area in figure 1. In addition, users are subject to different characteristics and/or behaviors even when they use the same amount of GP service. I am interested in whether changes in a particular explanatory variable in equation (5) changes the distribution of M along the quantiles of the vertical axis.

III.II Econometric Methodology: Quantile Regression Count Model (QRCM)

---

[1] Health utility index (HUI) is a generic, preference-scored, comprehensive system for measuring health status, health-related quality of life, and producing utility scores. HUI is calculated by the Health Utilities Group at McMaster University.

The commonly-used count data models including Poisson model, negative binomial model, zero-inflated model, TPM/hurdle model and LCM/finite mixture model are often criticized for their inflexibility on overdispersion. TPM and LCM[2] have become more popular in recent years because they outperform traditional Poisson and negative binomial models in allowing for heterogeneity among users. For instance, TPM assumes non-users and users behave differently. Based on this assumption, TPM estimates in two stages. The first part explains the binary choice of either zero utilization or non-zero utilization using logit or probit model. The second part explains the non-zero utilization using a truncated Poisson or negative binomial model. The TPM is appealing in that it provides the possibility of exploring heterogeneity between different utilization behaviors; however, most people are interested more in the information across the distribution of GP visits than in only at the mean level. QR, in this context, can be informative. QRCM can transform the discrete dependent variable to a continuous one, and then apply the standard QR on it to investigate the characteristics at different conditional quantiles of the dependent variable. On the one hand, if there is difference between quantiles, it is not enough to simply perform a TPM unless only the information at mean level is needed. On the other hand, if there is no difference between quantiles, it is unnecessary and redundant to perform TPM since the second stage will give the same result as in the first stage. In either case, QRCM has the potential to be a superior tool to TPM.

III.II.I. Quantile Regression (QR)

QR was proposed by Koenker and Basset (1978). The αth quantile for a random variable Y, which is characterized by a right continuous distribution function $F(y) = P(Y \leq y)$, is found by ordering observations of the variable Y from low to high, and locate the specific observation

[2] LCM is an approach more suitable for panel data. Cross-sectional data in this paper is not a good application of LCM.

with $\alpha \in (0,1)$ proportion of the observations below it and the remaining $1-\alpha$ proportion of observations above it. This specific observation is called the αth quantile of Y. We can express the above procedure by:

$$F^{-1}(\alpha) = \inf\{y : F(y) \geq \alpha\} \qquad (6)$$

Where $F^{-1}(\alpha)$ is called the αth quantile of Y.

Quantile regression (QR) models the relationship between X and the conditional quantiles of Y given X=x. Researchers are often interested in the two extremes of the conditional quantiles when there is heteroscedasticity in the OLS residuals[3]. QR is able to provide a more complete picture of the conditional distribution of Y given X=x when both lower and upper quantiles are of interest. QR is particularly useful when the rate of the change in the conditional quantile, expressed by the regression coefficients, depends on the quantile. QR has been increasingly popular in recent years in that it offers the flexibility of investigating the change in location and shape of the conditional distributions of Y at interested quantiles. The intuition of quantile regression lies on the existence of heterogeneous conditional distribution of Y.

However, QR deals with continuous dependent variable only. QR tries to minimize an objective function constructed by applying symmetric or asymmetric weight on the residuals of observations. Therefore, it requires the objective function to be continuous in order to perform Taylor expansion as we do on ordinary maximization or minimization problems. Applying QR directly on discrete count data therefore gives the estimation of conditional quantiles with

[3] White's test for heteroscedasticity investigates the possibility that the variance of the residuals from an OLS regression is a function of a set of explanatory variables, usually those which were included in the original regression. Quantile regression moves this analysis back a stage, analyzing the distribution of Y itself, rather than of the OLS residuals, and allows for more general changes in the shape of the conditional distribution of the dependent variable beyond its variance.

improper rate of convergence. It is not appropriate to apply QR before discrete data is smoothened.

III.II.II. Quantile Regression for Count Model (QRCM)

QRCM is proposed by Machado and Santos Silva (2005). The basic idea is to artificially smooth the data using a specific form of the jittering technique which was introduced by Stevens (1950). The major problem applying QR on a count data variable Y is that Y is discrete, thus $Q_Y(\alpha \mid X)$, the αth quantile of Y, cannot be a continuous function of covariates X, which makes it impossible to solve the linear programming problem used for QR. The purpose of jittering is to construct a continuous variable Z by creating a variable U, which is uniformly distributed in the interval [0,1) and independent of Y and X, and add it to Y, i.e., Z=Y+U. However, the distribution of Z may still not be smooth over the whole domain. Standard QR cannot be applied. To tackle this problem, a monotone transformation $T(Z;\alpha)$ is applied in order to satisfy the following restriction.

$$Q_{T(Z;\alpha)}(\alpha \mid X) = X^{'}\beta(\alpha) \tag{7}$$

Where $\hat{\beta}(\alpha)$ is the estimated coefficient vector of covariates at αth quantile.

Here at least a continuous explanatory variable exists in covariate X and $Q_T(\alpha \mid X)$ represents the αth quantile of the jittered variable $T_{Z;\alpha}$ and 0< α <1.

By applying the above procedure, $\Pr(T^{-1}(X^{'}\beta(\alpha)) = 0$ is ensured. Here $T^{-1}(\cdot)$ is defined as the inverse of the transformation $T(Z;\alpha)$. The monotone transformation $T(Z;\alpha)$ therefore guarantees at almost every realization of the vector of covariates x, the conditional density of z at the interested quantile will be continuous. Machado and Santos Silva specify the following monotone transformation before running the linear QR.

$$T(Z;\alpha)=\begin{cases}\log(Z-\alpha)\ for Z>\alpha\\ \log(\varsigma)\ for Z\le\alpha\end{cases} \qquad (8)$$

Where ς is a suitably small positive number[4].

The intuition of applying this monotone transformation is that quantiles are equivariant to monotonic transformation and invariant to censoring from below up to the quantile of interest. Based on the above transformation, $\beta(\alpha)$ can be estimated by running a standard linear QR of $T(Z;\alpha)$ on covariates x. The estimator is asymptotically normal and inference on the basis of t, LR, and Wald statistics is valid. Further on, the $100\alpha th$ quantile of Z and Y can be estimated by

$Q_z(\alpha\mid X)=\alpha+\exp(X^{'}\beta(\alpha))$ (9) and $Q_Y(\alpha\mid X)=\lceil Q_Z(\alpha\mid X)-1\rceil$

(10)

where $\lceil a\rceil$ denotes the ceiling function that returns the smallest integer greater than or equal to a.

Since U is introduced to smooth the discrete data and U itself is a noise term from a uniform distribution, it is natural to run Monte Carlo to average out the noise effect. The final estimate is therefore based on the average of the estimates from a certain number of jittered samples.

We know from the above description that there is a one-to-one relationship between the conditional quantiles of Z and Y. Because of the monotone transformation on Z, the relationship between coefficient estimates $\beta(\alpha)$ and Z or Y is essentially non-linear. It would be hard to interpret $\beta(\alpha)$ in terms of Z or Y. The signs on $\beta(\alpha)$ can give us an idea on the marginal effect of covariates on Z or Y. The Marginal Effect of a covariate on $Q_Z(\alpha\mid\bar{X})$ is often calculated, where $\bar{X}$ is a vector containing the mean value of the other covariates.

---

[4] Machado and Santos Silva (2005) use 1.0E-10. I use the same number in this paper.

IV. Model Description

IV.I Data

One challenge in research on cross-national comparison is the comparability of the data if they are from different resources. In this paper, I use data from Joint Canada/United States Survey of Health (JCUSH), 2002/3. JCUSH is jointly conducted by the Health Statistics Division of Statistics Canada and the National Center for Health Statistics (NCHS) of the U.S. The project is a one-time survey and the data was collected from 3,505 Canadians and 5,183 Americans who were eighteen or older, living in households with landline telephone from November 4, 2002 to March 31, 2003. 56 observations in number of consultations on GP, 249 in marital status variables, 12 in self-assessed health status, 332 in HUI, 276 in education dummies, 230 in immigration, 210 in physical activity index dummies, 39 in smoking dummies, 75 in number of chronic diseases and 348 in BMI dummies are dropped due to insufficient information from the respondents. The final sample size is 7721, with 4558 Canadian respondents and 3163 American respondents.

IV.II GP Utilization Model and Variable Specification

The GP utilization models for Canada and the U.S. are the following:

Canada Mode: $$y_i = \alpha + \sum_{j=1}^{p} x_{ij}\beta_p + cadrugins_i \beta_{cadrugins} + \varepsilon_i \qquad (11)$$

U.S. Model: $$y_i = \alpha + \sum_{j=1}^{p} x_{ij}\beta_p + medicare_i \beta_{medicare} + medicaid_i \beta_{medicaid}$$

$$+ nins_i \beta_{prvtins} + ucins_i \beta_{ucins} + \varepsilon_i \qquad (12)$$

where i=1,…,n and n=3163 for Canada model and 4558 for the U.S. model;

p=24: the number of common covariates;

Y=dependent variable: number of GP consultations;

$X = (X_1, X_2, ..., X_p)$: common covariate vector for both countries[5];

Country-specific covariates include *cadrugins* for Canada; and , *medicare*, *medicaid*, *nins* and *ucins* for the U.S.. In runs where the two countries samples are pooled, we include a public insurance variable, *pubins* [7], to represent Canadian Medicare.

Variable definitions are provided in table 1 and summary statistical description for the explanatory variables are shown in table 2. The explanatory variables are mostly dummies except for age, age2, household size, income, income square, health utility index, number of chronic diseases and number of chronic diseases squared. Among 7721 respondents, 4219 are females, 1256 are immigrants[8], 6380 have a regular doctor, 5152 have hospital insurance and 6657 respondents have good or better health. 15.08% Canadian respondents do not have a regular doctor while the number for American respondents is 18.96%. 2401 Canadian respondents have drug insurance. Among American respondents, 1019 have Medicare, 287 have Medicaid, 3407 have private insurance, 471 have no any health insurance and 158 did not tell what kind of health insurance they hold.

The distributions of number of consultations on GP for Canadians and Americans are provided in table 3 and Figures 2 & 3 respectively. A higher proportion of Canadian respondents uses more than four GP consultations than their American counterparts. Figure 4 shows the distribution of smokers in the two countries. There are more heavy smokers in Canada than in the U.S. More strikingly, there are around 10% more Americans than Canadians who never

---

[5] I do not include race dummies since race for black information is not available for Canadians. Test on the significance of race dummies on American sample shows race dummies are all insignificant.
[7] In Canada model, public insurance variable is not included because public insurance variable is 1 for all the Canadian respondents and it cannot be estimated in Canada model.
[8] JCUSH does not have the information on how many Canadian immigrants are Americans and how many American immigrants are Canadians.

smoked. The BMI distribution in two countries is presented in figure 5. A higher proportion of Canadians than Americans has normal weight. The proportion of overweight persons in the two countries is similar, but a higher proportion of Americans than Canadians are obese.. Income histograms for Canadians and Americans are shown in figures 6 & 7 respectively. The income is adjusted by GDP PPP 2003. Generally speaking, Americans have higher income than Canadians. If we look at the right tail of the histogram, we can see respondents with more than 110k total household income are all Americans. The proportion of Canadians or Americans who have or have no a regular doctor is shown in figure 8. Although all the Canadians have benefit from public health insurance, there are still around 15% Canadians do not have a regular doctor and who presumably go to a walk in clinic or emergency room when they are not well. The Americans number is fairly close to the Canadian: around 19% Americans do not have a regular doctor.

V. Econometric Result and Discussion

In this section, QRCM results are presented. Because TPM is increasingly popular in recent years in count data model analysis, I also present the TPM result and compare it with the QRCM result. Since the relationship between dependent variable and explanatory variables is not linear in either model, it is hard to interpret the coefficients of the explanatory variables. I thus provide the marginal effect of the covariates and compare them across different models and quantiles. A marginal effect can be interpreted as the increase on dependent variable when there is one unit increase on that specific variable keeping all other variables at their mean levels. In the discussion section, I discuss the results of QRCM for Canada and the U.S., compare the results of QRCM for the two countries, present the result from TPM and compare it with those from QRCM. I draw the graphs for marginal effects at different quantiles for key explanatory variables, i.e., sex, self-assessed health status, health utility index and having a regular doctor, from QRCM and Zero-Truncated Negative Binomial (ZTNB), which is the second part of the TPM, in figures 22. The graphs of marginal effects from QRCM and ZTNB for Candian drug insurance and immigrant in Canada are also provided in Figure 20 and 21. The dark, solid, thick horizontal lines in the graphs are the marginal effects from ZTNB and the dashed horizontal line means the marginal effect from TPM is insignificant (as for the immigrant case).

V.I Quantile Regression for Count Model Results

In quantile regression for count data model analysis, I am interested in the marginal effect of different explanatory variables across different quantiles and in whether the explanatory variables change the distribution of GP visits at different quantiles in and across Canada and the U.S. One obvious question is which quantiles to investigate. The unconditional distribution of GP visit provides the rough locations where zero, one-time, twice and high users locate. Since around 20% of the total respondents are zero GP users, and around 22% respondents use one GP consultation, I can roughly locate different users at different conditional quantiles. Taking the American sample as an example, at the 0.15 conditional quantiles, GP consultations is 0.87 with covariates at their mean levels. This means that at 0.15 quantile, there are both non-users and one-time users. At the 0.2 quantile, GP consultation is one, meaning most observations around 0.15 quantiles are one-time GP users. Between the 0.2 and 0.45 quantiles, there is a mixture of one- and two-time users. At 0.5 conditional quantile, most observations are two-time users. And at 0.7 conditional quantiles, most users are three-time users. Conditional quantiles 0.8 and 0.9 show the characteristics of high GP users. At quantile lower than 0.15, most respondents are zero users, it is pointless to investigate the structure of GP utilization for non-users. I run quantile count data model at 0.15, 0.2, 0.25, 0.4, 0.5, 0.6, 0.7, and 0.9 for Canada and the U.S. separately and report the marginal effect results in table 4 and 5.

V.I.I Canada Sample

In table 4, I report the marginal effect result for the Canada sample by running jittering 1,300 times. It is notable that across all the quantiles, sex, self-assessed health status, number of chronic diseases, having a regular doctor, and Canadian drug insurance are consistently significant at 5% significance level. Age, age square, health utility index, number of chronic diseases square, obese level II, and current non-smoker are mostly significant except at one or two quantiles. Among education variables, only lower than high school education is only significant at 0.15 quantile. All the other education variables are insignificant across all the quantiles. Among BMI variables, obese level II and III are significant at some quantiles and the left BMI variables are insignificant across all the quantiles. Income is everywhere insignificant. Income squared is significant at 0.7 quantile, while insignificant at the rest quantiles. Being an immigrant is only significant between the 0.4 and 0.6 quantiles.

For age, sex, self-assessed health status, health utility index, number of chronic disease, and having a regular doctor, the absolute scales of their marginal effects increase as quantile increases. Self-assessed health status and health utility index have negative marginal effect on the number of GP consultations across all the quantiles. Lower than high school education has negative marginal effect on the number of GP consultations at 0.15 quantile. The marginal effect of Canadian drug insurance stays fairly flat across different quantiles except after 0.7 quantile, where the marginal effect begins to increase. In addition, at 0.15 and 0.9 quantiles, less variables are significant than at middle quantiles.

V.I.II U.S. Sample

QRCM results for the U.S. sample are provided in table 5. Sex, self-assessed health status, health utility index, number of chronic diseases, number of chronic diseases square and having a regular doctor are consistently significant at all the quantiles. The Non-insured dummy is significant everywhere but at the 0.9 quantile. Age is significant between 0.25 and 0.5 quantiles and at 0.7 quantile. Age square is significant only between 0.25 and 0.45 quantiles. Among the education variables, only lower than high school education is significant, at the 0.15 quantile, and the remaining education variables are all insignificant across all quantiles. Income and income squared are everywhere insignificant. The BMI variables are significant on and off, while among the smoking variables, being a daily smoker is consistently insignificant, being an occasional smoker variable is significant between 0.15 and 0.25 quantiles and being a current non-smoker variable is only significant at the 0.8 and 0.9 quantiles. Being an Immigrant is only significant between the 0.4 and 0.5 quantiles. Medicare, Medicaid and the "unclear insurance status" variables are consistently insignificant.

Among those consistently significant variables, i.e., sex, self-assessed health status, health utility index, number of chronic diseases, number of chronic diseases and having a regular doctor, the scales of the marginal effects increase as quantile goes up. Self-assessed health status and health utility index have negative marginal effects on the number of GP consultations across all the quantiles. Lower than high school education has negative marginal effect on the number of GP consultations at 0.15 quantile. Non-insured respondents use less GP consultations than the respondents with either Medicare, or Medicaid, or private insurance. Current non-smokers use more GP consultations at high quantiles, and occasional smokers use more GP service at low quantiles.

V.I.III Discussion

One question that is often brought up in physician's service utilization literature is the endogeneity of health capital variables such as self-assessed health status and health utility index. In this paper, I tried the first stage IV regression on health utility index and used the predicted health utility index as an explanatory variable in the QRCM models. The results of using health utility index and predicted health utility index are quite similar. I therefore did not adjust health utility index and use it directly in the model. [9]

[9] I did not adjust self-assessed health status for the following reasons. Self-assessed health status is a 0/1 variable. Based on Angrist and Krueger (2001), it is sometimes more problematic to run a probit or a logit on an endogenous dummy variable. Although using a linear regression for the 1st stage IV estimates generates consistent second-stage estimates for an endogenous dummy variable, I could not find solid support on how to set the cut-off when predicting the dummy variable in the 1st stage IV estimates.

For the health utility index variable, I regress it on marital status variables—single, divorced and widow variables, physical activity index variables—active and moderate active variables, smokers variables—daily smoker, occasional smoker and current non-smoker variables, and reasons for unmet health care need variables—care not available, waiting time too long, cost too high and other reasons variables. The reasons for using the above variables are based on some preliminary result from QRCM when they are firstly included in the QRCM models. Physical activity index variables and marital status variables are consistently insignificant for both countries. The results from the 1st stage IV regression for both countries are as follows:

For Canada: adjusted R^2=0.1158

$$\begin{aligned} hui = {} & \underset{(0.008)}{0.886} - \underset{(0.009)}{0.001} single - \underset{(0.011)}{0.074} divr - \underset{(0.013)}{0.112} widow + \underset{(0.009)}{0.07} paia + \underset{(0.008)}{0.06} paim \\ & - \underset{(0.01)}{0.037} dsmoker + \underset{(0.016)}{0.006} osmoker - \underset{(0.008)}{0.019} cnsmoker - \underset{(0.023)}{0.117} carena \\ & - \underset{(0.02)}{0.076} wtlong - \underset{(0.043)}{0.212} costhigh - \underset{(0.016)}{0.161} otherreasons + residual \end{aligned}$$

For the U.S.: adjusted R^2=0.1248

$$\begin{aligned} hui = {} & \underset{(0.006)}{0.875} + \underset{(0.008)}{0.003} single - \underset{(0.009)}{0.046} divr - \underset{(0.01)}{0.119} widow + \underset{(0.008)}{0.089} paia + \underset{(0.008)}{0.068} paim \\ & - \underset{(0.009)}{0.034} dsmoker + \underset{(0.014)}{0.021} osmoker - \underset{(0.007)}{0.017} cnsmoker - \underset{(0.032)}{0.125} carena \\ & - \underset{(0.034)}{0.016} wtlong - \underset{(0.012)}{0.158} costhigh - \underset{(0.015)}{0.126} otherreasons + residual \end{aligned}$$

The numbers in the parenthesis are the standard errors of the relevant coefficient estimates. Adjusted R^2 for both countries are not high even though most covariates are significant. This is because health utility index, a proxy of health capital, comes from gross investment on health capital and the previous period's health capital. Since the cross-sectional data used in this paper cannot provide information on previous health capital, it is not possible to include the health capital variable from the last period in the 1st stage IV regression.

All the covariates in the above models are significant except for single, occasional smoker for both countries and waiting time too long for Canada. The negative signs on divorced and widow variable means those divorced and widowed respondents have lower health utility index, or their health situation is worse than married and single persons. The positive signs on physical activity index variables show that in comparison with those inactive persons, more active persons have better health even though they may not use more GP's service. Daily smokers and current

The result from QRCM is similar to those from previous literature in terms of the significance of key covariates. Sex, self-assessed health status, health utility index, having a regular doctor, number of chronic diseases and number of chronic diseases square are consistently significant across different quantiles. Self-assessed health status, health utility index, number of chronic diseases and number of chronic diseases square are health capital variables. Health capital is crucial in GP's service utilization and it has stronger impact on GP visits at high quantiles than at low quantiles.

The significant and negative marginal effects of self-assessed health status and health utility index in both countries could be interpreted as meaning that respondents with good health status use less GP consultations than respondents with poor health status (Figure 9). The distribution of the number of GP visits for respondents with good or better health shifts to the left relative to the distribution of the number of GP visits for respondents with poor or worse health, and the shifts are greater at high quantiles than at low quantiles. Thus there is not only a shift of the distribution, but also a change in the shape of the distribution of the number of GP consultations. The opposite case applies to sex and having a regular doctor in both countries (Figure 9). The distribution of the number of GP consultations for females shifts to the right relative to the distribution of the number of GP consultations for males. Females use more GP service than males in both countries, and a smaller proportion of females use zero or one GP consultations that males, keeping other variables constant. The distribution of the number of GP visits for respondents having a regular doctor shifts to the right of the distribution of the number of GP visits for respondents without a regular doctor. Again, the shapes of the distribution of the

---

non-smokers all have poorer health than non-smokers and occasional smokers. Those who complained they have unmet health care needs are of poorer health than those who did not complain, and may use more GP's service than those who did not report unmet needs.

number of GP visits are different for respondents with a regular doctor and those without a regular doctor.

When we look at the marginal effects of sex, self-assessed health status and health utility index (Figures 10-12), it is obvious that the magnitudes of the marginal effects in Canada are bigger than those in the U.S. Having a regular doctor tells a story about the availability of a family doctor and it is more associated with availability of GP service. Having a regular doctor can save patients a lot of time in waiting at a walk-in clinic or an emergency room. Interestingly, the marginal effects at different quantiles are almost the same in the two countries (Figure 13).

In Figure 14 and 15, which are the marginal effects of lower than high school education, we can see that it exhibits similar pattern in the two countries, i.e., it is only significant at 0.15 quantile. If we look at the distributions of the number of GP visits for lower than high school population and the rest population, the distribution of lower than high school population shifts to the left of the distribution of the rest population only at 0.15 quantile and stays the same as the rest population at quantiles higher than 0.15. This means that the lower than high school population are less likely to visit a GP for the first time than the remaining population. However, once they have visited GP for the first time, their behavior is not different from the rest of the population. Based on this, we could expect the marginal effect of lower than high school education population on the number of GP consultations from the first part of the TPM to be significant and negative, and that from the second part of the TPM to be insignificant. QRCM could then accommodate TPM in providing more information on how the distribution of the dependent variable changes.

For insurance variables, it is clear that Canadians with drug insurance use more GP consultations than Canadians without drug insurance. The marginal effect of Canadian drug insurance is fairly flat across different quantiles except at 0.9 quantile (Figure 16). In Canada, drug expenditure is the only out of pocket a Canadian faces if she/he does not have drug insurance and is not a senior. It is natural that Canadians with drug insurance tend to use more GP's service than those without drug insurance. Having no drug insurance seems to deter Canadians from seeing GPs, especially when they have to use more GP service. In the U.S., Medicare and Medicaid are consistently insignificant (Figure 17 and 18). However, if we look at the marginal effect from non-insured variable, we can see that Americans without any insurance coverage use less GP consultations than those with either Medicare, Medicaid or private insurance except at 0.9 quantile (Figure 19), where most respondents are very sick and insurance coverage may not be important any more; in addition, the difference in GP utilization between the non-insured group and the insured group increases as quantile increases. The results from insurance variables are consistent with the previous literature suggesting that people with insurance coverage use more physician service.

For the smoker variables, current non-smokers use more GP service than the rest population including daily smokers, non-smokers and occasional smokers. The reason might be those current non-smokers quit smoking because they have got health problems and were asked to quit smoking, while the negative effect from smoking for those daily smokers has not appeared yet. Those daily smokers investigated in the survey still enjoy smoking and their behavior appears to be the same as non-smokers and occasional smokers. For the U.S. case, smoker variables exhibit a more complicated pattern. However, the behavior of daily smokers and non-smokers remains the same, which is consistent with the result in Canada.

BMI tells if the respondent is underweight, normal weight, overweight or obese. BMI beyond normal range might be an accumulation of poor human capital across time. BMI variables exhibit complicated marginal effect on the number of GP visits in the two countries. In Canada, obese people use more GP service at mid-ranged quantiles, while in the U.S., in comparison with normal weight respondents, all the left population uses more GP service at different quantiles. Since individual's BMI may be associated with genetic characteristics, it partly explains why BMI variables often give mixed result at different quantiles.

According to Grossman's model, income influences respondent's budget constraint, and further influences respondent's choices on purchasing medical goods. Theoretically, most low users who are zero or one time users are usually in better health than high users. On the one hand, higher income would not make low-users to visit GP more since it would not bring them more utility by visiting GP more if their health status is good. On the other hand, increased income would make high-users able to purchase more medical goods. We would expect income to have positive effect at high quantiles and no impact at low quantiles. The American sample does not tell this story even though income appears to be weakly significant occassionally. Only in Canadian sample, income square is found to be significant at 0.7 quantile. This result shows income does not actually influence GP utilization. Deb and Trivedi (1997; 2002) tell similar stories. One possible reason is high income persons are mostly healthier than low income persons.

Immigrants use more GP service at the mid-ranged quantiles in both countries. Being an Immigrant is significant at more quantiles in Canada than in the U.S. and the marginal effect of immigrant on GP visits is bigger in Canada than in the U.S. The Canadian sample has higher percentage of immigrants than the U.S. Moreover, 150 out of 579 Canadian immigrants in the JCUSH are older than 65 years. Senior immigrants usually use the free health care service more than younger immigrants in Canada. In contrast, in the U.S., only 89 out of 677 American immigrants are older than 65 years. Immigrants do not automatically obtain free health insurance coverage; therefore, they buy insurance by themselves, obtain supplemental insurances from government or private company or stay uninsured. It is mostly the institutional difference in health systems in two countries that causes the difference in immigrant's impact in the two countries.

Age and number of chronic diseases enter the model with quadratic forms. I list the turning point for age and number of chronic diseases in table 6. The signs on age and age square mean that the marginal effect of age decreases as age increases, and after the turning point, the marginal effect of age increases as age increases. The turning point for age in the Canada sample increases as quantile goes up, and the case for the U.S. sample is similar even though age and age square are only significant at some of the quantiles in the middle. The signs on number of chronic diseases and number of chronic diseases square show that the marginal effect of number of chronic diseases increases first and then after the turning point, the marginal effect of number of chronic diseases decreases. The turning point for both countries is somewhere between 3 and 4 chronic diseases. This is a reasonable result since 97% of the population in the two countries has 3 or less than 3 chronic diseases. Therefore, the marginal effect of number of chronic diseases increases as the number of chronic diseases increases in both countries according to the survey information.

V.II Two-part Model result and comparison with the QRCM results

V.II.I TPM result

TPM is an increasingly popular count data model in applied work, so it is interesting to compare the results from QRCM and TPM and see which approach can offer us more information. I report TPM results and compare them with QRCM results in this section. I employ the logit model for the first part of TPM and ZTNB for the second part of TPM. Results for Canada sample and the U.S. sample are reported in table 7.

In the first part of TPM for two countries, sex, number of chronic diseases and having a regular doctor are consistently significant with positive impact on the probability of visiting a GP. Health utility index has consistently negative impact on the probability of seeing a GP. Self-assessed health status is insignificant for Canada sample, however, significant at 1% significance level for the U.S. sample. Lower than high school education and high school education dummies are negatively significant for Canada sample and insignificant for the U.S. sample. This implies that low-educated persons in Canada are less likely to visit GP than the rest respondents. Age, age square, obese class I and class II, and current non-smokers are significant for Canada sample, while insignificant for the U.S. sample. Underweight population significantly use less GP's service in the U.S., while use the same GP's service as the rest population in Canada. For country-specific insurance variables, Canadian drug insurance is significant at 1% significance level. In comparison with population with private insurance, only non-insured population use significantly less GP's service. This confirms the conclusion from former literature that Canadians or Americans with insurance coverage are more likely to visit GP. Among BMI variables, obese level II and III are significant in the Canada sample, while only under weight variable is significant in the U.S. sample.

In the second part ZTNB model, age, age square, sex, current non-smokers, self-assessed health status, health utility index, number of chronic diseases, number of chronic diseases square and having a regular doctor are consistently significant for two countries and the signs of the marginal effects of the above variables are the same as in the 1st part model. In addition, the scales of the marginal effects are much bigger than in the 1st part model. In comparison with respondents with normal weight respondents, obese level II respondents significantly use more GP's service among non-zero users in Canada and the U.S. Among smoker variables, current non-smokers and daily smokers all use significantly more GP's service than non-smokers in the U.S. In Canada, only currently non-smokers use significantly more GP's service than the rest population. Income is insignificant in both countries. For country-specific insurance variables, Canadians with drug insurance use more GP's service than those without drug insurance. In the U.S., insurance variables are all significant except for those respondents who did not reveal their insurance information; they use more GP's service than the rest population.

V.II.II Comparison with the QRCM results

Since TPM gives results based on logit and ZTNB models and QRCM gives results across quantiles, the coefficients from TPM and QRCM are not directly comparable. TPM and QRCM, however, define marginal effect the same way. It is therefore reasonable to compare the marginal effects of the same variable from the two models. The merit of QRCM lies on its ability to explore information across different conditional quantiles of GP visits. ZTNB, the second part of TPM, gives marginal effect estimation of the positive GP visits at mean level. By comparing the marginal effect from ZTNB and QRCM (Figures 20-22), we can see QRCM provides more information at different conditional distribution of the number of GP visits, while ZTNB provides information at the mean level of the positive GP visits.

For the U.S. sample, key variables, i.e., age, age square, sex, self-assessed health status, health utility index, number of chronic diseases, number of chronic diseases square and having a regular doctor, are similar in both models in terms of significance. Age is significant in ZTNB, however, however, only partly insignificant in QRCM for Americans. Immigrant does not appear to be significant in TPM, however, significant at mid-quantiles in QRCM in both countries.

Daily smoker is only significant for American users in ZTNB, while never significant in QRCM for both Canadians and Americans. The most notable difference between two models is on non-insured variable. Non-insured variable appears to be significant in the logit model but insignificant in ZTNB model for American respondents. The Non-insured variable is on and off at different quantiles in QRCM. Medicare and Medicaid remains insignificant in both models. QRCM and ZTNB models give mixed results on non-insured variable and consistent result on Medicare and Medicaid. Other than that, TPM result is quite consistent with the result from QRCM.

For the Canada sample, the results of ZTNB show the marginal effects of age, age square, sex, income, income square, self-assessed health status, health utility index, number of chronic diseases, number of chronic diseases square, having a regular doctor, and Canadian drug insurance are consistent in QRCM and ZTNB in terms of significance. ZTNB model only uses around 80% of the observations because zero users are truncated. The mean level of 80% of the whole sample locate at around 0.6-0.7 quantile of the whole sample. In Figures 20-22, the marginal effect estimates from ZTNB are at around marginal effect at 0.5-0.7 quantile of the QRCM except for having a regular doctor variable, which is consistent with the above analysis. The QRCM marginal effect estimates for key explanatory variables, such as sex, self-assessed health status, health utility index and having a regular doctor, show the distribution of the number of GP visits shifts more at high quantiles than at low quantiles conditional on the explanatory variables and this is obviously not available from ZTNB models. ZTNB model obviously cannot provide us with more information about the heterscedastisity of the dependent variable than QRCM does.

VI. Conclusion

The result in this paper shows QRCM is appealing in that its estimates provide more information on how the distribution of the dependent variable changes at different conditional quantiles than do some other estimation methods. At the very least, QRCM appears to be a good supplement to parametric count models such as TPM. The difference in the distribution of GP utilization conditional on the covariates shows it is necessary to perform QRCM. The QRCM results show sex, self-assessed health status, health utility index, having a regular doctor and number of chronic diseases are important factors influencing the distribution of the number of GP visits. The impact from the above important factors increases from low conditional quantiles to high conditional quantiles in both countries. The key variables shifts the location of the distribution of the dependent variable more at high quantiles than at low quantiles, and more in Canada than in the U.S. except for having a regular doctor variable. Insurance increases the GP's service utilization at different quantiles in both countries and the impact is bigger at high quantiles than at low quantiles. Other variables including age, immigrant, smoker variables and BMI variables are also important sometimes, although not as robust as the above-mentioned key variables. Income does not show significant influence on GP's service utilization. The major structure difference in GP utilization between two countries lies on the scale of the marginal effects from important covariates. The scales of the marginal effects from the important covariates in Canada are generally higher than those in the U.S except for having a regular doctor, which has similar impact on the GP's service utilization in the two countries.

Future research of this paper will be focusing on performing formal statistical test on the difference between the determinants of the GP's service utilization in the two countries. Moreover, other health care utilization topics such as specialist and hospital utilization may also be investigated.

References:

Alfonso Miranda, (2006) "QCOUNT: Stata program to fit quantile regression models for count data", Statistical Software Components S456714, Boston College Department of Economics.

Angrist, J.D. and Krueger, A.B. (2001) "Instrumental variables and the search for identification: from supply and demand to natural experiments", The Journal of Economic Perspectives, Vol. 15, No. 4: 69-85.

Armstrong, H., Clement, W., Lin, Z. and Prus, S. (2006) "Contrasting Inequalities: Comparing correlates of health in Canada and the United States" Social and Economics Dimensions of An Aging Population (SEDAP) working paper No. 167 (available at http://ideas.repec.org/p/mcm/sedapp/167.html).

Atella, V., Brindisi, F., Deb, P., Rosati, F.C. (2004) "Determinants of access to physician services in Italy: a latent class seemingly unrelated probit approach", Health Economics 13: 657-668

Deb, P. and Trivedi, P.K. (1997) "Demand for medical care by the elderly: a finite mixture approach", Journal of Applied Econometrics, 12: 313-336.

Deb, P. and Trivedi, P.K. (2002) "The structure of demand for health care: latent class versus two-part models", Journal of Health Economics, 21: 601-625.

Denavas-Walt, C., Proctor, B.D., and Smith, J. (2007) "Income, poverty and health insurance coverage in the United States: 2006", U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau.

Dunlop, S., Coyte, P.C. and McIsaac, W. (2000) "Socio-economic status and the utilization of physicians' services: results from the Canadian National Population Health Survey" Social Science & Medicine 51: 123-133.

Grossman M. (1972a) "On the concept of health capital and the demand for health", Journal of Political Economy, 80: 223-235.

Grossman M. (1972b) "The demand for health: A theoretical and empirical investigation", NBER, Columbia University Press: New York.

Grossman, M. (1999), "The Human Capital Model of the Demand for Health", Working Paper 7078 prepared for the Handbook of Health Economics.

Jimenez-Martin, S., Labeaga, J.M. and Martinez-Granado, M. (2004) "An empirical analysis of the demand for physician services across the European Union", European Journal of Health Economics, 5: 150-165.

Koenker, R. and Bassett, (1978), "Regression Quantiles", Econemetrica 46: 33-50.

Lahiri, K., Xing, G. (2004) "An econometric analysis of veterans' health care utilization using two-part models", Empirical Economics, 29: 431-449.

Lasser, K.E., Himmelstein, D.U. and Woolhandler, S. (2006) "Access to care, health status, and health disparities in the United States and Canada: results of a cross-national population-based survey", American Journal of Public Health, July, 96, 7: 1300-1307.

Machado, J. A. F. and Santos Silva, J.M.C. (2005) "Quantiles for Counts", Journal of the American Statistical Association, 100, 472: 1226-1237.

Rodriguez, M. and Stoyanova, A. (2004) "The effect of private insurance access on the choice of GP/specialist and public/private provider in Spain", Health Economics 13: 689-703.

Sanmartin, C., Berthelot, J-M, Ng, E., Murphy, K., Blackwell, D.L., Gentleman, J.F., Martinez, M.E. and Simlle, C.M. (2006) “Comparing health and health care use in Canada and the United States”, Health Affairs, July, 25: 1122-1142.

Sarma, S. and Simpson, W. (2006) “A microeconometric analysis of Canadian health care utilization”, Health Economics 15: 219-239.

Stabile M. (2001) “Private insurance subsidies and public health care markets: evidence from Canada”, Canadian Journal of Economics 34: 921-942.

Stevens, W.L. (1950), “Fiducial limits of the Parameter of a Discontinuous distribution,” Biometrika, 37, 117-129.

Table 1: Variable Definition

cons_gp	number of consulations to family doc./GP in 12 months
age	(age of respondent)/10
age2	age^2
sex	female=1 and male=0
hhlds	household size, household size=5 means 5 or more persons in the hhld
immigrant	country of birth of Canadian (American) respondents is not Canada (the U.S.)=1, otherwise=0
edulhs	highest edu. attained-- less than high school=1, otherwise=0
eduhs	highest edu. attained--high school degree or equivalent=1, otherwise=0
educ	highest edu. attained—trades cert. voc. sch./comm.. col./cegep=1, otherwise=0
eduu#	highest edu. attained—university or coll. cert. incl. below. bach.=1, otherwise=0
income	total household income (10,000 USD PPP), those who did not answer this question get 0
income2	income^2
nsincomed	respondents who did not answer total household income question=1,otherwise=0
bmiu	(BMI<18.5)=1, otherwise=0
bmin#	(18.5<=<BMI<25)=1, otherwise=0
bmiow	(25<=BMI<30)=1, otherwise=0
bmiob1	obese class I (30<= BMI<=34.9)=1, otherwise=0
bmiob2	obese class II (35<= BMI<=39.9)=1, otherwise=0
bmiob3	obese class III (BMI>=40)=1, otherwise=0
dsmoker#	daily smoker=1, otherwise=0
osmoker	occasional smoker (former daily for >3 mths & not former daily for >3 mths)=1, otherwise=0
cnsmoker	current non-smoker (smoked or not smoked >100 cig. in life)=1,otherwise=0
nvsmoker	never smoked 100 or a whole cig.=1, otherwise=0
sahs	self-assessed health status is good or better =1, otherwise, sahs=0
hui	health utility index
nchrncd	number of chronic diseases (asthma, arthritis, high blood pressure, emphysema or chronic obstructive pulmonary disease, diabetes, heart disease, angina and heart attack)
nchrncd2	nchrncd^2
hvrd	have a regular doctor=1, otherwise=0
cadrugins	Canadians having insurance for prescription medication=1,otherwise=0
pubins;ca	Canadians having public health insurance=1, otherwise=0, otherwise=0 ; ca is Canada dummy which is exactly the same as pubins.
nins	Americans having no health insurance coverage=1, otherwise=0
medicare	Americans having Medicare coverage=1, otherwise=0
medicaid	Americans having Medicaid coverage=1, otherwise=0
prvtins#	Americans having private insurance health coverage=1, otherwise=0
ucins	Americans having unclear health insurance coverage=1, otherwise=0
Variables used in the first stage IV (dsmoker, osmoker, cnsmoker and nvsmoker are also used):	
married#	married=1, otherwise=0
single	single=1, otherwise=0
divr	separated or divorced=1, otherwise=0
widow	widow=1 if respondent is widowed, otherwise=0
paia	physical activity index: active=1, otherwise=0
paim	physical activity index: moderate=1, otherwise=0
paii#	physical activity index: inactive=1, otherwise=0

wtlong	waiting time too long=1, otherwise=0
costhigh	cost too high=1, otherwise=0
otherreasons	reasons other than carena, wtlong and costhigh=1, otherwise=0

variables with # are used as reference groups and dropped to avoid multicolinearity.

Table 2: Statistical Description (N=7721)

Variable	U.S. (n=4558)				Canada (n=3163)			
	Mean	s.d	Min	Max	Mean	s.d	Min	Max
age	4.8284	1.7123	1.8	8.5	4.7358	1.7318	1.8	8.5
age2	26.2447	17.5821	3.24	72.25	25.4256	17.6637	3.24	72.25
sex	0.5586	0.4966	0	1	0.5289	0.4992	0	1
hhlds	2.1924	1.2617	1	5	2.4224	1.2137	1	5
immigrant	0.1485	0.3557	0	1	0.1831	0.3868	0	1
edulhs	0.1099	0.3128	0	1	0.2137	0.4100	0	1
eduhs	0.3649	0.4814	0	1	0.2858	0.4519	0	1
educ	0.1444	0.3515	0	1	0.2197	0.4141	0	1
eduu	0.3809	0.4857	0	1	0.2807	0.4494	0	1
income	3.7147	3.8011	0	13	3.5722	3.2344	0	10.92
income2	28.2445	43.5716	0	169	23.2185	32.5077	0	119.2464
nsincomed	0.3080	0.4617	0	1	0.2409	0.4277	0	1
bmiu	0.0228	0.1493	0	1	0.0288	0.1672	0	1
bmin	0.4296	0.4951	0	1	0.4774	0.4996	0	1
bmiow	0.3357	0.4723	0	1	0.3396	0.4736	0	1
bmiob1	0.1413	0.3484	0	1	0.1157	0.3199	0	1
bmiob2	0.0467	0.2111	0	1	0.0266	0.1608	0	1
bmiob3	0.0239	0.1528	0	1	0.0120	0.1090	0	1
dsmoker	0.1661	0.3722	0	1	0.1989	0.3992	0	1
osmoker	0.0546	0.2273	0	1	0.0588	0.2353	0	1
cnsmoker	0.3894	0.4877	0	1	0.4448	0.4970	0	1
nvsmoker	0.3899	0.4878	0	1	0.2975	0.4572	0	1
sahs	0.8559	0.3513	0	1	0.8713	0.3349	0	1
hui	0.8587	0.2190	-0.257	1	0.8696	0.2081	-0.243	1
nchronicd	0.6957	1.0714	0	7	0.5880	0.9899	0	7
nchronicd2	1.6316	4.0700	0	49	1.3253	3.7418	0	49
hvrd	0.8104	0.3920	0	1	0.8492	0.3579	0	1
cadrugins					0.7591	0.4277	0	1
medicare	0.2236	0.4167	0	1				
medicaid	0.0630	0.2429	0	1				
prvtins	0.7475	0.4345	0	1				
ucins	0.0347	0.1829	0	1				
married	0.5608	0.4963	0	1	0.5827	0.4932	0	1
single	0.1867	0.3897	0	1	0.2106	0.4078	0	1
divr	0.1505	0.3576	0	1	0.1192	0.3241	0	1
widow	0.1020	0.3027	0	1	0.0876	0.2827	0	1
paia	0.2102	0.4075	0	1	0.2611	0.4393	0	1
paim	0.2095	0.4070	0	1	0.2637	0.4407	0	1
paii	0.5803	0.4936	0	1	0.4752	0.4995	0	1
carena	0.0094	0.0967	0	1	0.0253	0.1570	0	1
wtlong	0.0081	0.0897	0	1	0.0338	0.1808	0	1
costhigh	0.0676	0.2510	0	1	0.0066	0.0812	0	1
otherreasons	0.0410	0.1984	0	1	0.0487	0.2152	0	1

Table 3: Frequency distribution of the dependent variable

No. of consultations on GP	CA			U.S.		
	freq.	cum. Per.	proportion	freq.	cum. Per.	proportion
0	591	18.68	0.187	951	20.86	0.209
1	685	40.34	0.217	1,058	44.08	0.232
2	569	58.33	0.180	930	64.48	0.204
3	344	69.21	0.109	499	75.43	0.109
4	320	79.32	0.101	416	84.55	0.091
5	144	83.88	0.046	132	87.45	0.029
6	142	88.37	0.045	196	91.75	0.043
7	25	89.16	0.008	29	92.39	0.006
8	38	90.36	0.012	49	93.46	0.011
9	9	90.64	0.003	19	93.88	0.004
10	38	91.84	0.012	58	95.15	0.013
11	5	92	0.002	1	95.17	0.000
12	155	96.9	0.049	115	97.7	0.025
13	6	97.09	0.002	2	97.74	0.000
14	4	97.22	0.001	5	97.85	0.001
15	17	97.76	0.005	24	98.38	0.005
16	2	97.82	0.001	6	98.51	0.001
17	1	97.85	0.000	2	98.55	0.000
18	2	97.91	0.001	6	98.68	0.001
20	28	98.8	0.009	22	99.17	0.005
21	1	98.83	0.000	1	99.19	0.000
24	11	99.18	0.003	9	99.39	0.002
25	2	99.24	0.001	4	99.47	0.001
26	2	99.3	0.001	4	99.56	0.001
30	4	99.43	0.001	4	99.65	0.001
31	18	100	0.006	16	100	0.004
Total	3,163			4,558		

Table 4: Quantile Regression Count Model Result----CA

Quantile	0.15	0.2	0.25	0.4	0.45	0.5	0.6	0.7	0.9
age	-0.088	-0.164***	-0.215**	-0.305*	-0.422*	-0.461*	-0.737*	-0.928*	-2.447*
	(0.093)	(0.099)	(0.101)	(0.101)	(0.119)	(0.141)	(0.186)	(0.23)	(0.566)
age2	0.011	0.017***	0.021**	0.028*	0.037*	0.04*	0.062*	0.076*	0.189*
	(0.009)	(0.01)	(0.01)	(0.01)	(0.011)	(0.013)	(0.017)	(0.023)	(0.06)
sex	0.295*	0.363*	0.416*	0.541*	0.533*	0.571*	0.608*	0.8*	1.736*
	(0.051)	(0.056)	(0.06)	(0.065)	(0.075)	(0.083)	(0.117)	(0.14)	(0.305)
hhlds	0.01	-0.016	0.012	-0.001	-0.012	0.002	-0.025	-0.033	0.199
	(0.024)	(0.025)	(0.026)	(0.03)	(0.036)	(0.036)	(0.051)	(0.065)	(0.13)
immigrant	0.081	0.127	0.157***	0.29*	0.283*	0.337*	0.318**	0.224	-0.16
	(0.064)	(0.08)	(0.083)	(0.099)	(0.098)	(0.105)	(0.131)	(0.169)	(0.403)
edulhs	-0.196*	-0.134***	-0.124	-0.023	0.02	0.023	0.092	0.108	0.409
	(0.069)	(0.081)	(0.082)	(0.107)	(0.119)	(0.121)	(0.182)	(0.196)	(0.543)
eduhs	-0.109***	-0.047	-0.028	-0.039	-0.035	-0.102	-0.05	-0.001	0.009
	(0.059)	(0.065)	(0.07)	(0.081)	(0.093)	(0.103)	(0.155)	(0.187)	(0.452)
educ	-0.012	0.002	0.006	-0.033	-0.028	-0.026	0.089	0.044	-0.022
	(0.066)	(0.069)	(0.071)	(0.081)	(0.092)	(0.113)	(0.156)	(0.183)	(0.489)
income	-0.015	0.007	-0.009	-0.039	-0.048	-0.113	-0.089	-0.178***	-0.248
	(0.041)	(0.042)	(0.044)	(0.048)	(0.061)	(0.069)	(0.08)	(0.107)	(0.222)
income2	0.001	0	0	0.004	0.005	0.01***	0.008	0.017**	0.013
	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.006)	(0.006)	(0.008)	(0.018)
nsincomed	-0.048	-0.011	-0.126	-0.15	-0.167	-0.377**	-0.296	-0.383	-0.274
	(0.11)	(0.121)	(0.122)	(0.122)	(0.159)	(0.166)	(0.212)	(0.302)	(0.632)
bmiu	-0.032	-0.146	-0.161	0.214	0.22	0.2	0.133	0.855	0.873
	(0.098)	(0.102)	(0.141)	(0.252)	(0.292)	(0.297)	(0.246)	(0.646)	(1.117)
bmiow	-0.021	0.017	0.104	0.118	0.078	0.072	0.074	0.028	0.91***
	(0.054)	(0.06)	(0.065)	(0.072)	(0.078)	(0.084)	(0.125)	(0.147)	(0.526)
bmiob1	-0.014	0.005	0.047	0.032	0.021	-0.007	0.104	0.103	0.025
	(0.082)	(0.088)	(0.098)	(0.104)	(0.109)	(0.147)	(0.184)	(0.215)	(0.398)
bmiob2	0.461***	0.639	0.97*	1.023*	1.054*	1*	1.69*	1.854***	4.416
	(0.249)	(0.47)	(0.236)	(0.174)	(0.31)	(0.298)	(0.486)	(1.044)	(2.797)
bmiob3	0.164	0.412	0.758**	0.83***	0.823	1.104***	0.806**	0.867	1.339
	(0.44)	(0.532)	(0.373)	(0.504)	(0.626)	(0.594)	(0.369)	(0.982)	(1.393)
dsmoker	-0.079	-0.065	-0.05	-0.047	0.007	0.133	0.254	0.185	-0.287
	(0.077)	(0.083)	(0.094)	(0.1)	(0.119)	(0.128)	(0.182)	(0.22)	(0.455)
osmoker	0.168	0.12	0.386**	0.432*	0.445*	0.454**	0.601**	0.215	-0.182
	(0.166)	(0.149)	(0.153)	(0.165)	(0.17)	(0.203)	(0.264)	(0.31)	(0.587)
cnsmoker	0.162*	0.124**	0.141**	0.144**	0.192**	0.238*	0.352*	0.324***	0.297
	(0.058)	(0.063)	(0.063)	(0.073)	(0.08)	(0.091)	(0.124)	(0.175)	(0.399)
sahs	-0.326*	-0.489*	-0.518*	-0.805*	-0.791*	-0.866*	-1.18*	-1.501*	-2.951*
	(0.12)	(0.163)	(0.164)	(0.151)	(0.16)	(0.164)	(0.255)	(0.465)	(0.497)
hui	-0.451***	-0.683**	-1.16*	-1.477*	-1.746*	-2.252*	-3.156*	-4.4*	-9.86*
	(0.236)	(0.278)	(0.295)	(0.288)	(0.325)	(0.541)	(0.971)	(0.988)	(2.052)
nchronicd	0.363*	0.403*	0.478*	0.572*	0.697*	0.776*	0.991*	1.256*	2.437*
	(0.05)	(0.054)	(0.06)	(0.069)	(0.083)	(0.083)	(0.105)	(0.17)	(0.507)
nchronicd2	-0.044*	-0.053*	-0.061*	-0.071*	-0.095*	-0.119*	-0.155*	-0.19*	-0.378***
	(0.014)	(0.014)	(0.013)	(0.019)	(0.02)	(0.02)	(0.025)	(0.05)	(0.194)
hvrd	0.759*	0.917*	1.159*	1.385*	1.539*	1.7*	1.932*	2.064*	3.296*
	(0.043)	(0.047)	(0.053)	(0.06)	(0.069)	(0.075)	(0.158)	(0.167)	(0.346)

Quantile	0.15	0.2	0.25	0.4	0.45	0.5	0.6	0.7	0.9
cadrugins	0.195*	0.209*	0.335*	0.348*	0.268*	0.277*	0.333*	0.373**	0.873**
	(0.054)	(0.061)	(0.064)	(0.076)	(0.085)	(0.094)	(0.122)	(0.147)	(0.388)
y=Qz(α |X)	0.943	1.153	1.544	1.917	2.117	2.313	2.896	3.644	6.60356

1. Figures in parenthesis are standard errors. The symbols *, **, and *** denote significant at 1, 5, and 10% significance levels respectively.
2. To get the number of jittered samples, I start from 1000 samples and increase the number by 100 until there are no significant changes in the estimation. The final jittered samples are 1300 for Canada and 1200 for the U.S.
3. I am grateful to the qcount package in stata provided by Alfonso Miranda (2006).
4. Please refer to table 1 for the variable definitions.

Table5: Quantile Regression Count Model Result----U.S.

Quantile	0.15	0.2	0.25	0.4	0.45	0.5	0.6	0.7	0.9
Variable	M.E.	M.E.	M.E.	M.E.	M.E.	M.E.	M.E.	M.E.	M.E.
age	-0.034	-0.027	-0.143**	-0.177**	-0.227*	-0.213**	-0.129	-0.297**	-0.937***
	(0.065)	(0.063)	(0.064)	(0.071)	(0.076)	(0.085)	(0.115)	(0.137)	(0.509)
age2	0.004	0.003	0.014**	0.015**	0.018**	0.016***	0.008	0.024***	0.059
	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	(0.012)	(0.014)	(0.055)
sex	0.121*	0.117*	0.136*	0.209*	0.197*	0.231*	0.293*	0.301*	0.46***
	(0.033)	(0.034)	(0.035)	(0.042)	(0.043)	(0.047)	(0.062)	(0.077)	(0.256)
hhlds	-0.003	0.004	-0.002	-0.006	-0.003	0.006	0.024	0.013	0.127
	(0.013)	(0.013)	(0.014)	(0.017)	(0.018)	(0.02)	(0.028)	(0.033)	(0.114)
immigrant	0.023	0.051	0.066	0.14**	0.181**	0.151**	0.105	0.148	0.375
	(0.05)	(0.054)	(0.053)	(0.068)	(0.071)	(0.074)	(0.088)	(0.125)	(0.328)
edulhs	-0.172*	-0.101***	-0.052	-0.079	-0.042	0.002	-0.006	0.049	-0.495
	(0.051)	(0.058)	(0.069)	(0.083)	(0.088)	(0.109)	(0.113)	(0.135)	(0.434)
eduhs	-0.048	-0.015	0.02	-0.006	-0.019	0.066	0.124***	0.028	0.09
	(0.037)	(0.036)	(0.039)	(0.044)	(0.047)	(0.055)	(0.073)	(0.09)	(0.282)
educ	-0.025	-0.011	0.018	0.051	0.029	0.064	0.071	0.009	0.36
	(0.047)	(0.046)	(0.049)	(0.058)	(0.058)	(0.062)	(0.084)	(0.104)	(0.294)
income	0.032	0.017	0.02	0.022	0.026	0.024	-0.002	-0.012	-0.064
	(0.022)	(0.021)	(0.022)	(0.025)	(0.027)	(0.033)	(0.043)	(0.053)	(0.167)
income2	-0.002	-0.001	-0.002	-0.003	-0.003***	-0.003	-0.001	-0.002	0.002
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.013)
nsincomed	0.038	0.008	0.004	-0.034	-0.052	-0.08	-0.219***	-0.28***	-0.629
	(0.07)	(0.069)	(0.07)	(0.08)	(0.087)	(0.101)	(0.13)	(0.156)	(0.528)
bmiu	0.204**	0.191***	0.316**	0.354**	0.334**	0.349**	0.316	0.227	1.313**
	(0.093)	(0.099)	(0.145)	(0.147)	(0.169)	(0.175)	(0.222)	(0.457)	(0.665)
bmiow	0.076***	0.082**	0.118*	0.182*	0.174*	0.181*	0.159**	0.087	-0.125
	(0.041)	(0.04)	(0.042)	(0.047)	(0.049)	(0.054)	(0.071)	(0.091)	(0.272)
bmiob1	0.161*	0.148*	0.197*	0.236*	0.279*	0.324*	0.374*	0.229**	0.158
	(0.054)	(0.051)	(0.057)	(0.067)	(0.075)	(0.087)	(0.105)	(0.108)	(0.363)
bmiob2	0.055	0.082	0.09	0.135	0.17	0.324***	0.481*	0.647*	1.936
	(0.102)	(0.072)	(0.092)	(0.109)	(0.141)	(0.171)	(0.185)	(0.236)	(2.334)
bmiob3	0.164***	0.142***	0.173**	0.175	0.179	0.315	0.356***	0.563**	-0.203
	(0.088)	(0.086)	(0.087)	(0.144)	(0.151)	(0.229)	(0.208)	(0.28)	(0.57)
dsmoker	-0.036	-0.01	-0.041	-0.025	0.012	0.023	0.004	0.091	0.621
	(0.053)	(0.047)	(0.048)	(0.057)	(0.06)	(0.067)	(0.086)	(0.144)	(0.465)

Quantile	0.15	0.2	0.25	0.4	0.45	0.5	0.6	0.7	0.9
osmoker	0.152**	0.14**	0.165**	0.047	0.045	0.03	-0.007	0.121	0.143
	(0.076)	(0.069)	(0.069)	(0.089)	(0.09)	(0.112)	(0.164)	(0.189)	(0.415)
cnsmoker	-0.03	-0.024	-0.03	0.069	0.079***	0.084	0.109	0.167**	0.738*
	(0.036)	(0.035)	(0.037)	(0.043)	(0.045)	(0.052)	(0.067)	(0.085)	(0.26)
sahs	-0.195*	-0.225*	-0.28*	-0.372*	-0.435*	-0.458*	-0.65*	-1.092*	-1.962*
	(0.06)	(0.066)	(0.068)	(0.08)	(0.088)	(0.093)	(0.145)	(0.212)	(0.538)
hui	-0.564*	-0.706*	-0.661*	-1.094*	-1.217*	-1.386*	-1.731*	-2.436*	-6.729*
	(0.129)	(0.14)	(0.138)	(0.173)	(0.164)	(0.225)	(0.323)	(0.442)	(1.045)
nchronicd	0.267*	0.289*	0.327*	0.416*	0.443*	0.494*	0.597*	0.758*	1.564*
	(0.033)	(0.031)	(0.033)	(0.034)	(0.036)	(0.044)	(0.064)	(0.078)	(0.236)
nchronicd2	-0.035*	-0.039*	-0.043*	-0.05*	-0.055*	-0.066*	-0.086*	-0.111*	-0.256*
	(0.01)	(0.008)	(0.01)	(0.009)	(0.008)	(0.011)	(0.019)	(0.022)	(0.061)
hvrd	0.896*	1.054*	1.148*	1.446*	1.556*	1.655*	1.887*	1.984*	3.151*
	(0.033)	(0.032)	(0.033)	(0.041)	(0.043)	(0.048)	(0.063)	(0.1)	(0.265)
medicare	-0.002	-0.013	-0.03	-0.058	-0.052	-0.031	-0.018	-0.234***	0.528
	(0.055)	(0.053)	(0.052)	(0.067)	(0.07)	(0.078)	(0.105)	(0.132)	(0.567)
medicaid	0.075	0.017	0.023	0.019	0.039	0.036	0.172	0.338	0.853
	(0.063)	(0.063)	(0.066)	(0.095)	(0.09)	(0.104)	(0.169)	(0.208)	(0.691)
nins	-0.212*	-0.204*	-0.23*	-0.292*	-0.349*	-0.415*	-0.527*	-0.677*	-0.259
	(0.061)	(0.069)	(0.073)	(0.088)	(0.092)	(0.098)	(0.104)	(0.167)	(0.518)
ucins	0.125	0.097	0.129	0.201***	0.236***	0.121	0.295	0.329	1.48
	(0.099)	(0.096)	(0.104)	(0.12)	(0.131)	(0.145)	(0.189)	(0.232)	(1.092)
y=Qz(α |X)	0.831	1.001	1.190	1.640	1.793	1.967	2.336	2.982	5.453

*Please refer to table 1 for the variable definitions.

Table 6: Turning point for age and number of chronic diseases

variable	quantile	0.15	0.2	0.25	0.4	0.45	0.5	0.6	0.7	0.9
age	CA	40	48.2	50.7	54.5	57.0	57.6	59.4	61.1	64.7
	U.S.	42.5	45	51.1	59.0	63.1	66.6	80.6	61.9	79.4
number of chronic diseases	CA	4.1	3.8	3.9	4.0	3.7	3.3	3.2	3.3	3.2
	U.S.	3.8	3.7	3.8	4.2	4.0	3.7	3.5	3.4	3.1

Table 7: Two-Part Model result--M.E.

Variable	CA		U.S.	
	1st part	2nd part	1st part	2nd part
age	-0.067*	-0.884*	-0.041***	-0.51*
	(0.021)	(0.218)	(0.022)	(0.168)
age2	0.007*	0.066*	0.004***	0.04**
	(0.002)	(0.021)	(0.002)	(0.017)
sex	0.081*	0.659*	0.036*	0.258*
	(0.013)	(0.129)	(0.012)	(0.093)
hhlds	0.001	0.025	0.001	0.023
	(0.005)	(0.06)	(0.004)	(0.039)
immigrant	0.018	0.121	0.015	0.231
	(0.015)	(0.171)	(0.015)	(0.146)
edulhs	-0.046**	0.164	-0.036	-0.164
	(0.023)	(0.206)	(0.025)	(0.162)
eduhs	-0.031***	0.117	0.002	-0.15
	(0.017)	(0.178)	(0.013)	(0.108)
educ	-0.009	0.063	0.003	0.012
	(0.017)	(0.186)	(0.017)	(0.14)
income	0.004	-0.171***	0.008	-0.004
	(0.01)	(0.099)	(0.008)	(0.062)
income2	0	0.012	0	-0.003
	(0.001)	(0.008)	(0.001)	(0.004)
nsincomed	0	-0.335	0.002	-0.198
	(0.027)	(0.25)	(0.023)	(0.18)
bmiu	0.013	0.852***	0.088*	0.444
	(0.033)	(0.477)	(0.022)	(0.334)
bmiow	0.009	0.283***	0.008	0.009
	(0.013)	(0.15)	(0.012)	(0.107)
bmiob1	0.005	0.054	0.021	0.113
	(0.019)	(0.211)	(0.017)	(0.143)
bmiob2	0.062**	2.079*	-0.013	0.648**
	(0.029)	(0.613)	(0.031)	(0.263)
bmiob3	0.09*	0.822	0.06***	0.163
	(0.032)	(0.663)	(0.033)	(0.291)
dsmoker	0.005	0.041	-0.006	0.453*
	(0.017)	(0.198)	(0.017)	(0.158)
osmoker	0.037***	0.003	0.005	-0.024
	(0.02)	(0.292)	(0.023)	(0.211)
cnsmoker	0.03**	0.319**	-0.002	0.415*
	(0.014)	(0.153)	(0.013)	(0.108)
sahs	-0.022	-1.544*	-0.053*	-0.98*
	(0.024)	(0.28)	(0.018)	(0.184)
hui	-0.094**	-2.58*	-0.091**	-1.994*
	(0.044)	(0.331)	(0.037)	(0.224)
nchronicd	0.094*	1.066*	0.095*	0.821*
	(0.024)	(0.147)	(0.016)	(0.098)
nchronicd2	-0.004	-0.142*	-0.012*	-0.099*
	(0.01)	(0.034)	(0.005)	(0.021)
hvrd	0.317*	0.979*	0.411*	0.905*

	(0.027)	(0.176)	(0.021)	(0.121)
cadrugins	0.049*	0.308**		
	(0.015)	(0.146)		
medicare			0.007	-0.143
			(0.025)	(0.16)
medicaid			0.026	0.218
			(0.025)	(0.198)
nins			-0.077*	-0.081
			(0.023)	(0.192)
ucins			0.023	0.913*
			(0.026)	(0.341)
ttl # of obs.	3163	2572	4558	3607

*Please refer to table 1 for the variable definitions.

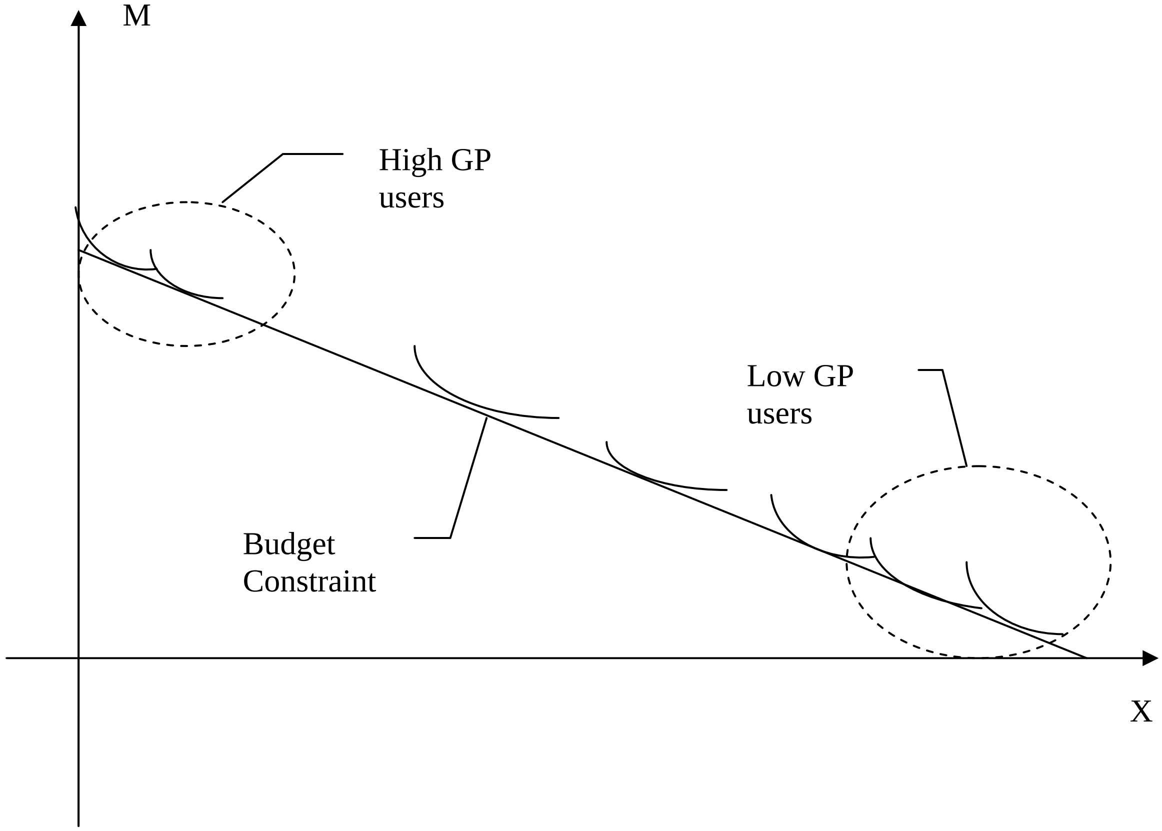

Figure 1: Grossman-model-based Health Care Utilization Model

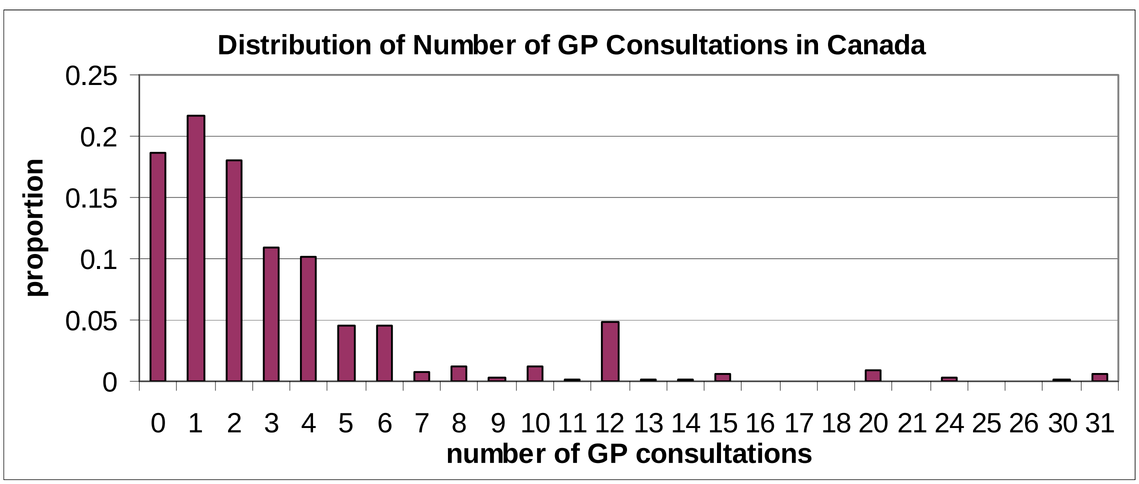

Figure 2

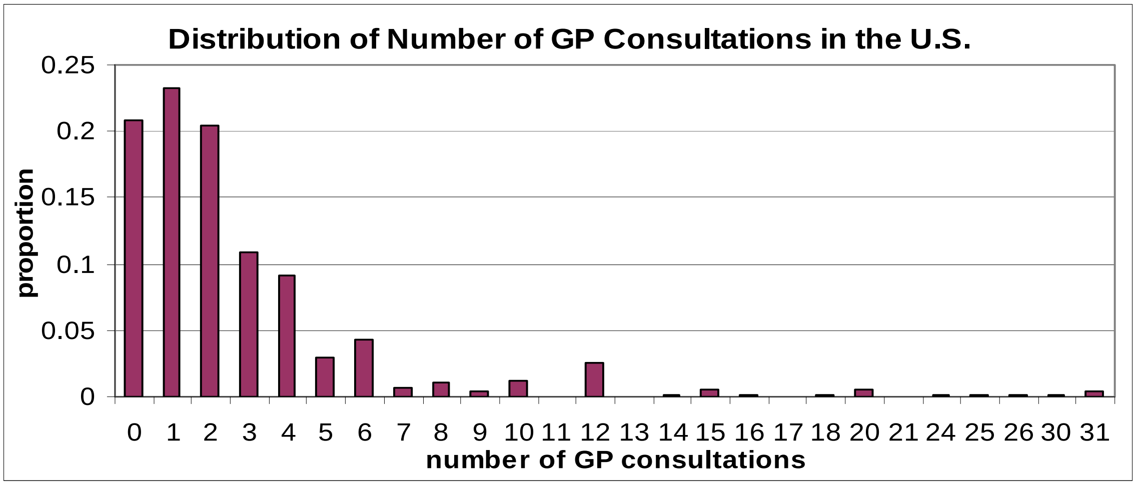

Figure 3

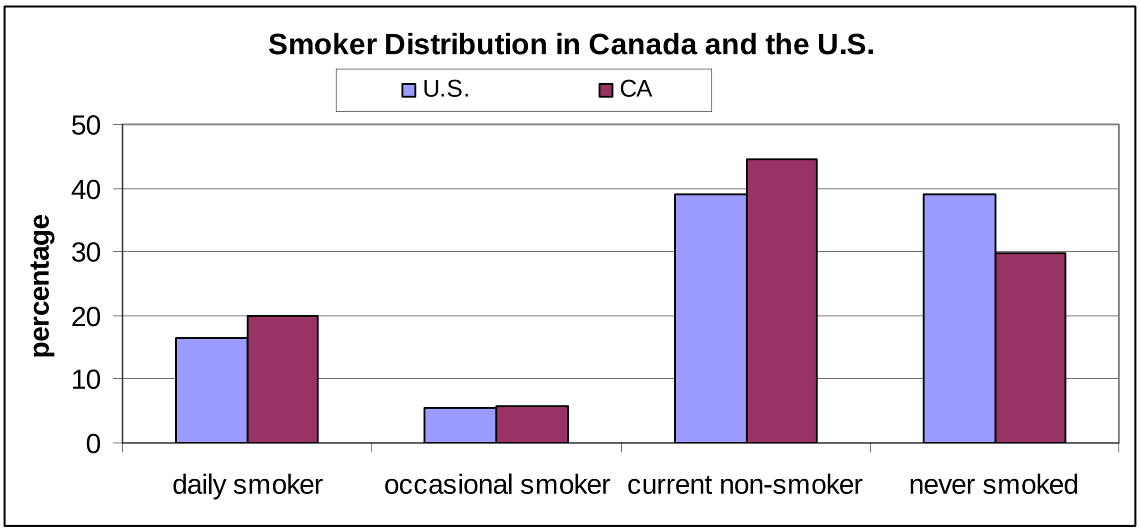

Figure 4

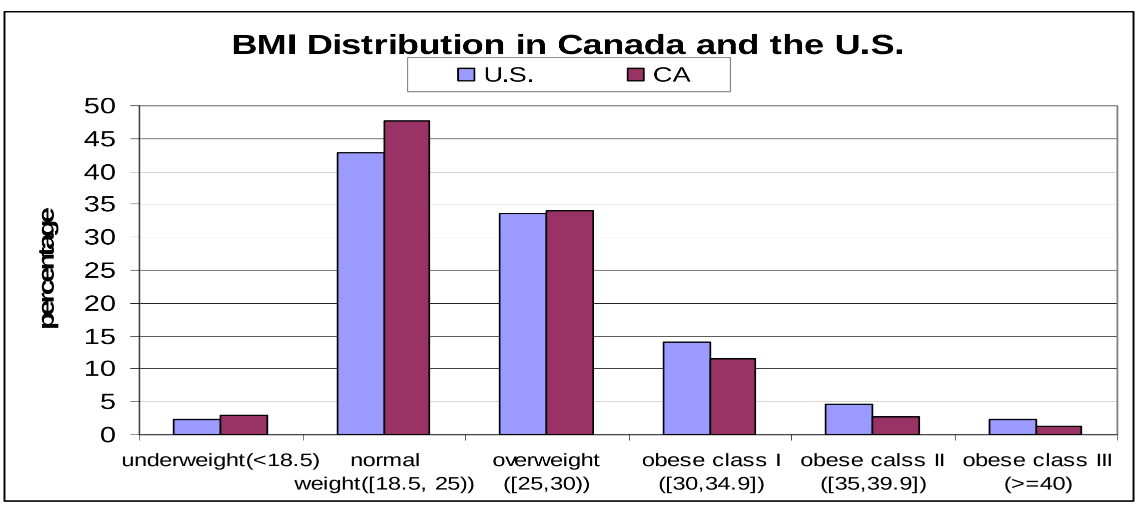

Figure 5

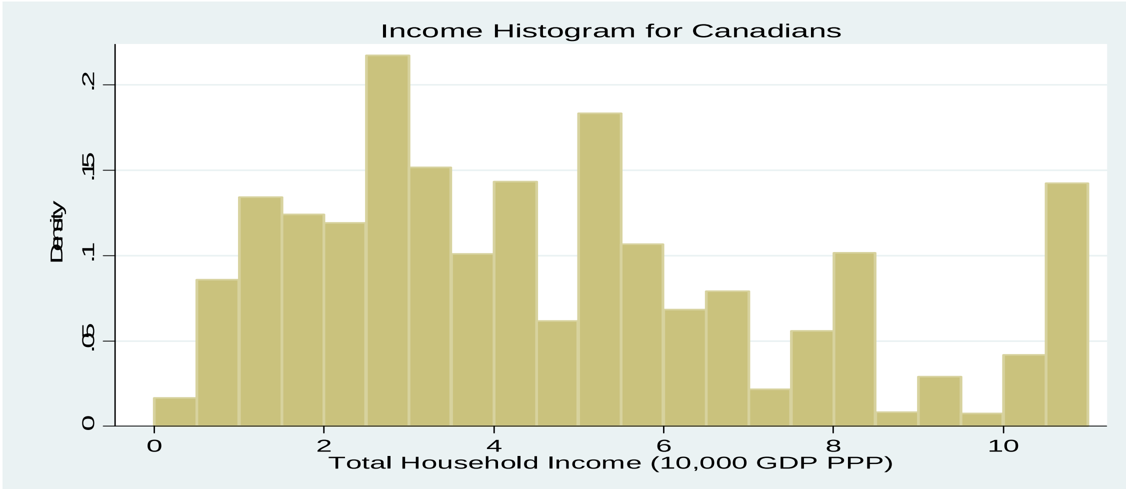

Figure 6

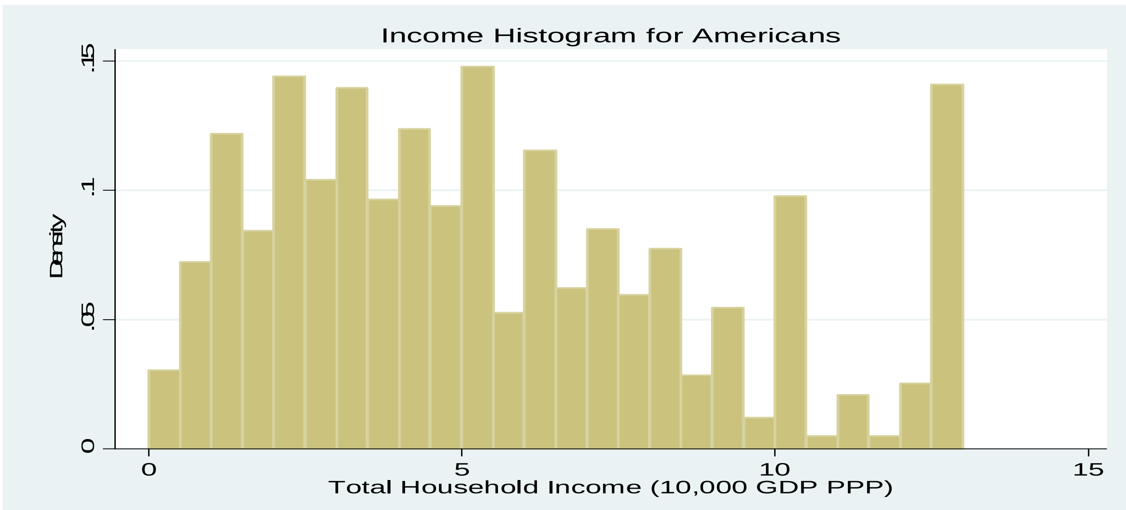

Figure 7

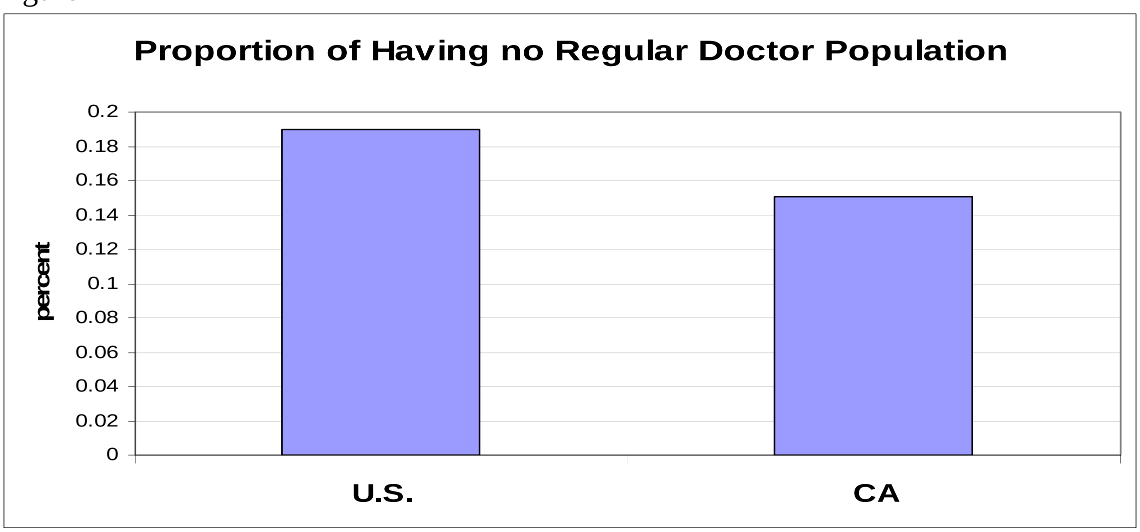

Figure 8

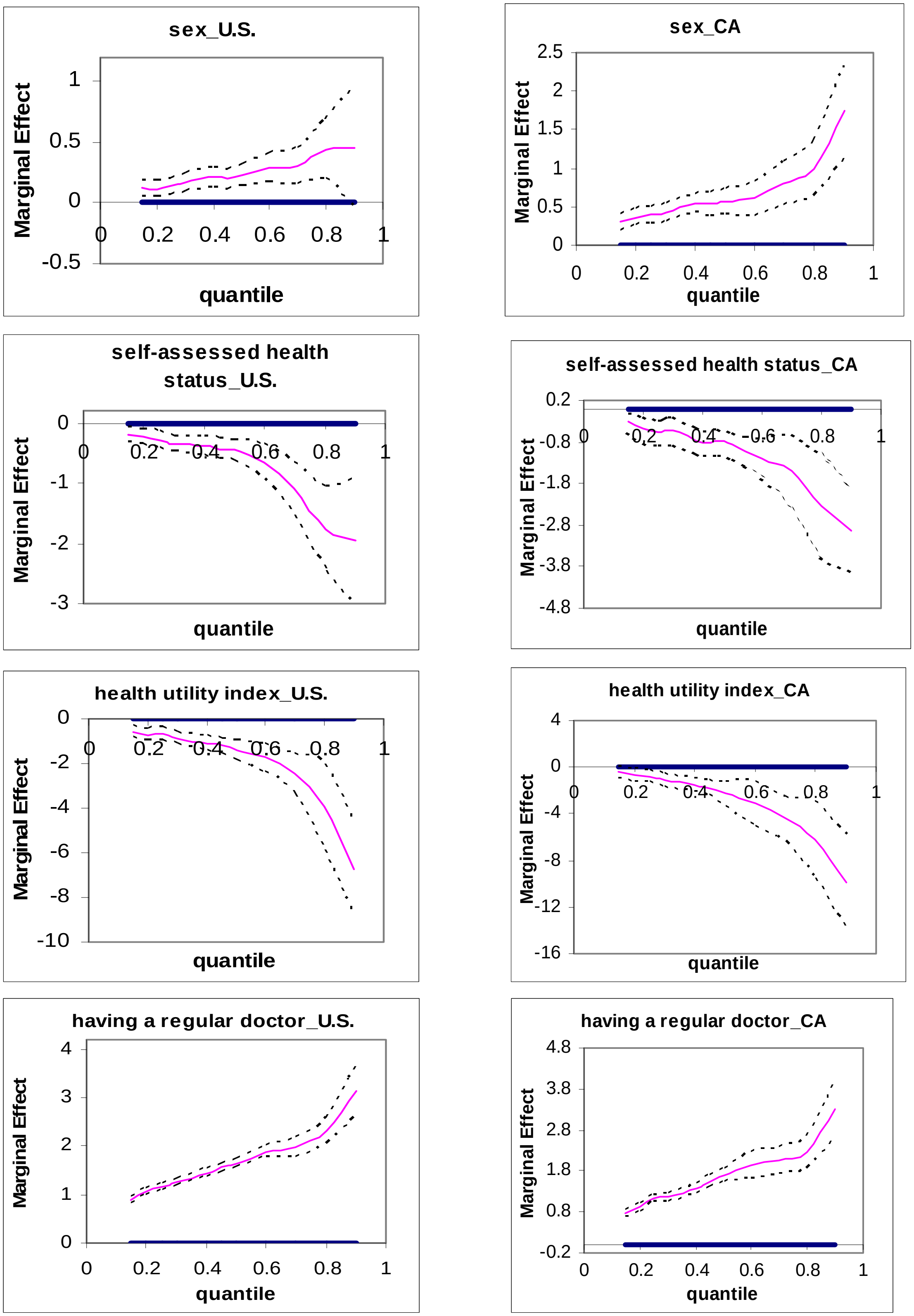
sex_U.S.
Marginal Effect
1
0.5
0
-0.5
0
0.2
0.4
0.6
0.8
1
quantile
sex_CA
Marginal Effect
2.5
2
1.5
1
0.5
0
0
0.2
0.4
0.6
0.8
1
quantile
self-assessed health status_U.S.
Marginal Effect
0
-1
-2
-3
0
0.2
0.4
0.6
0.8
1
quantile
self-assessed health status_CA
Marginal Effect
0.2
-0.8
-1.8
-2.8
-3.8
-4.8
0
0.2
0.4
0.6
0.8
1
quantile
health utility index_U.S.
Marginal Effect
0
-2
-4
-6
-8
-10
0
0.2
0.4
0.6
0.8
1
quantile
health utility index_CA
Marginal Effect
4
0
-4
-8
-12
-16
0
0.2
0.4
0.6
0.8
1
quantile
having a regular doctor_U.S.
Marginal Effect
4
3
2
1
0
0
0.2
0.4
0.6
0.8
1
quantile
having a regular doctor_CA
Marginal Effect
4.8
3.8
2.8
1.8
0.8
-0.2
0
0.2
0.4
0.6
0.8
1
quantile

Figure 9

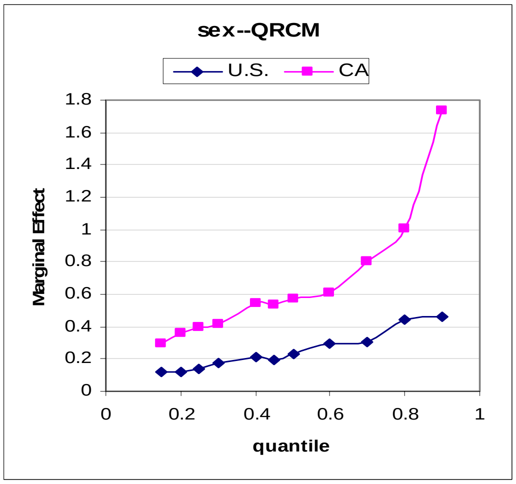

Figure 10

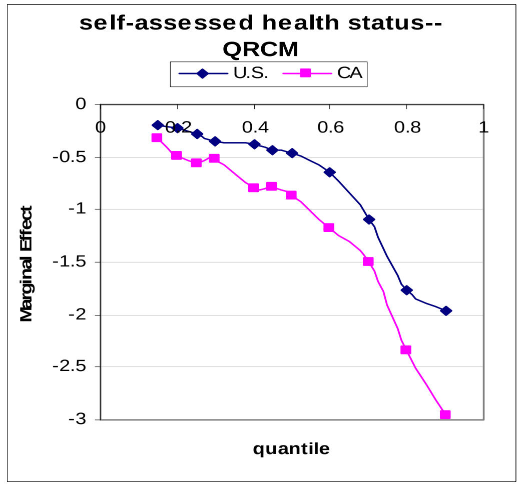

Figure 11

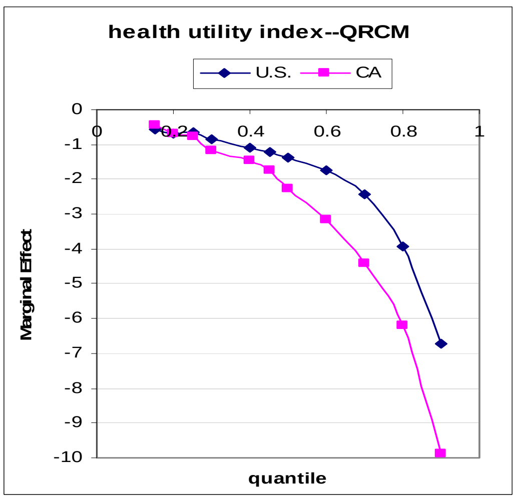

Figure 12

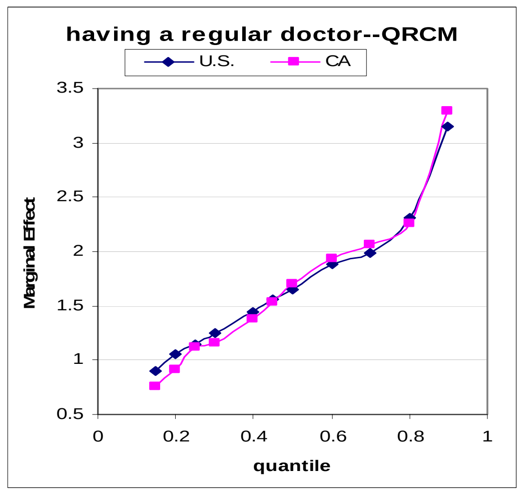

Figure 13

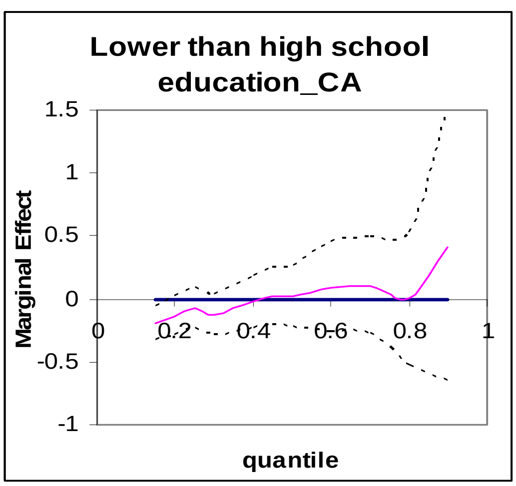

Figure 14

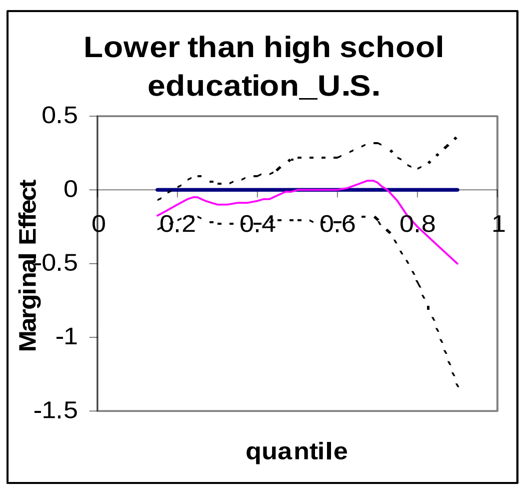

Figure 15

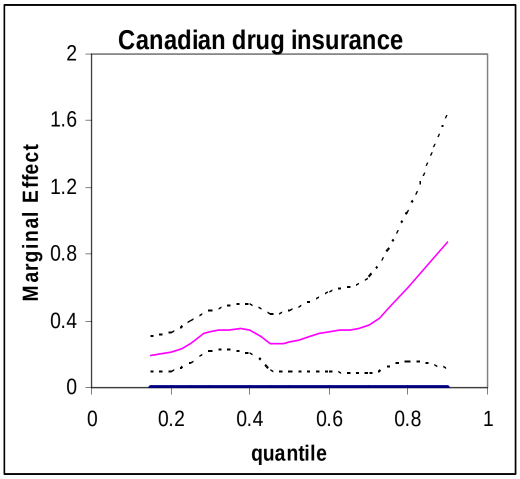

Figure 16

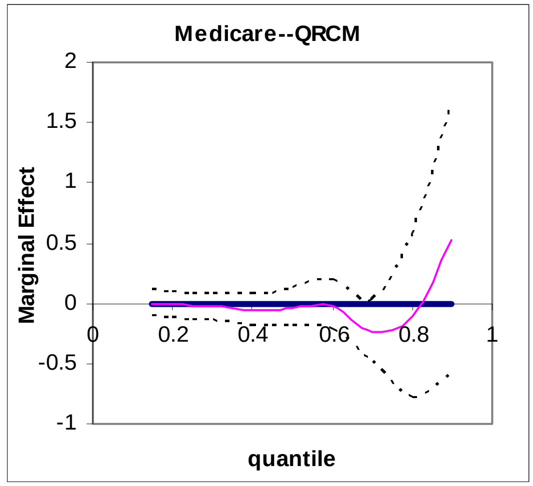

Figure 17

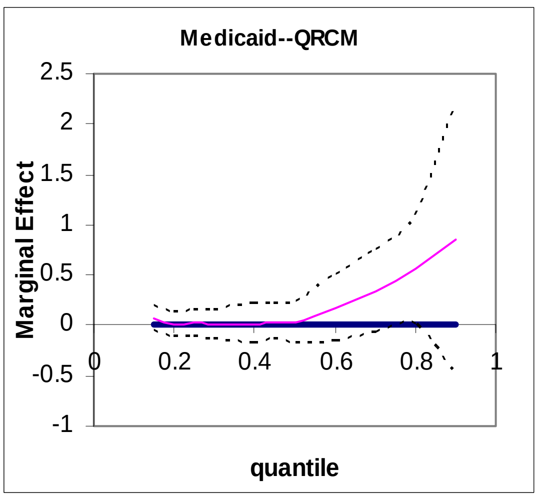

Figure 18

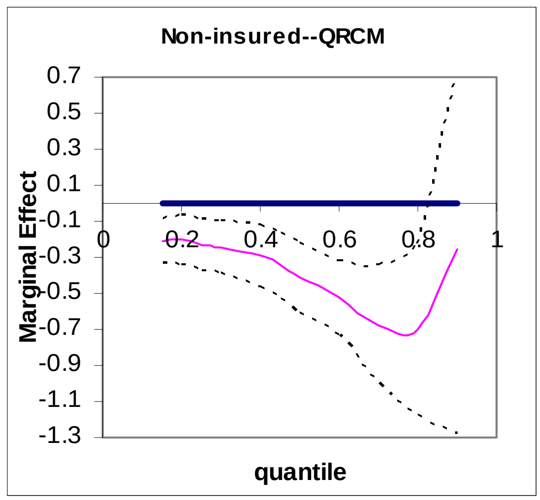

Figure 19

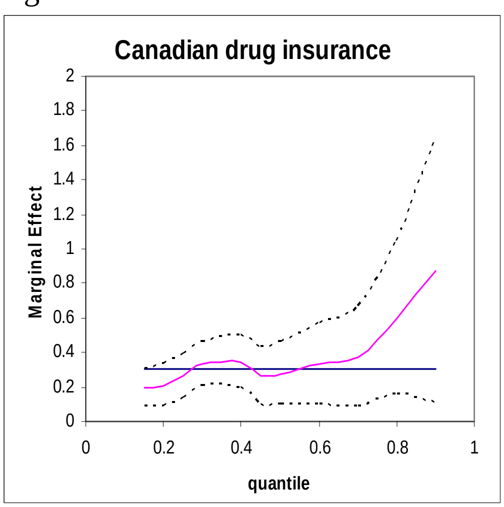

Figure 20

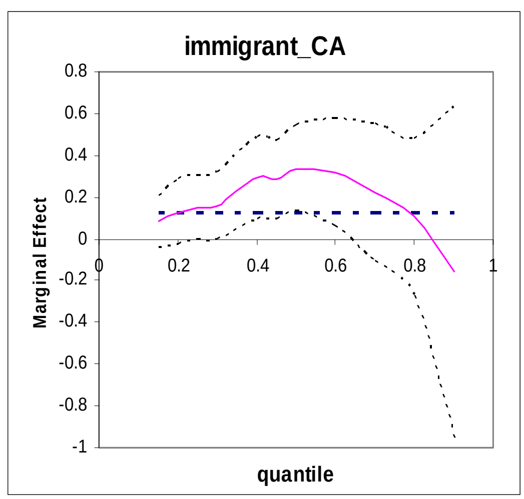

Figure 21

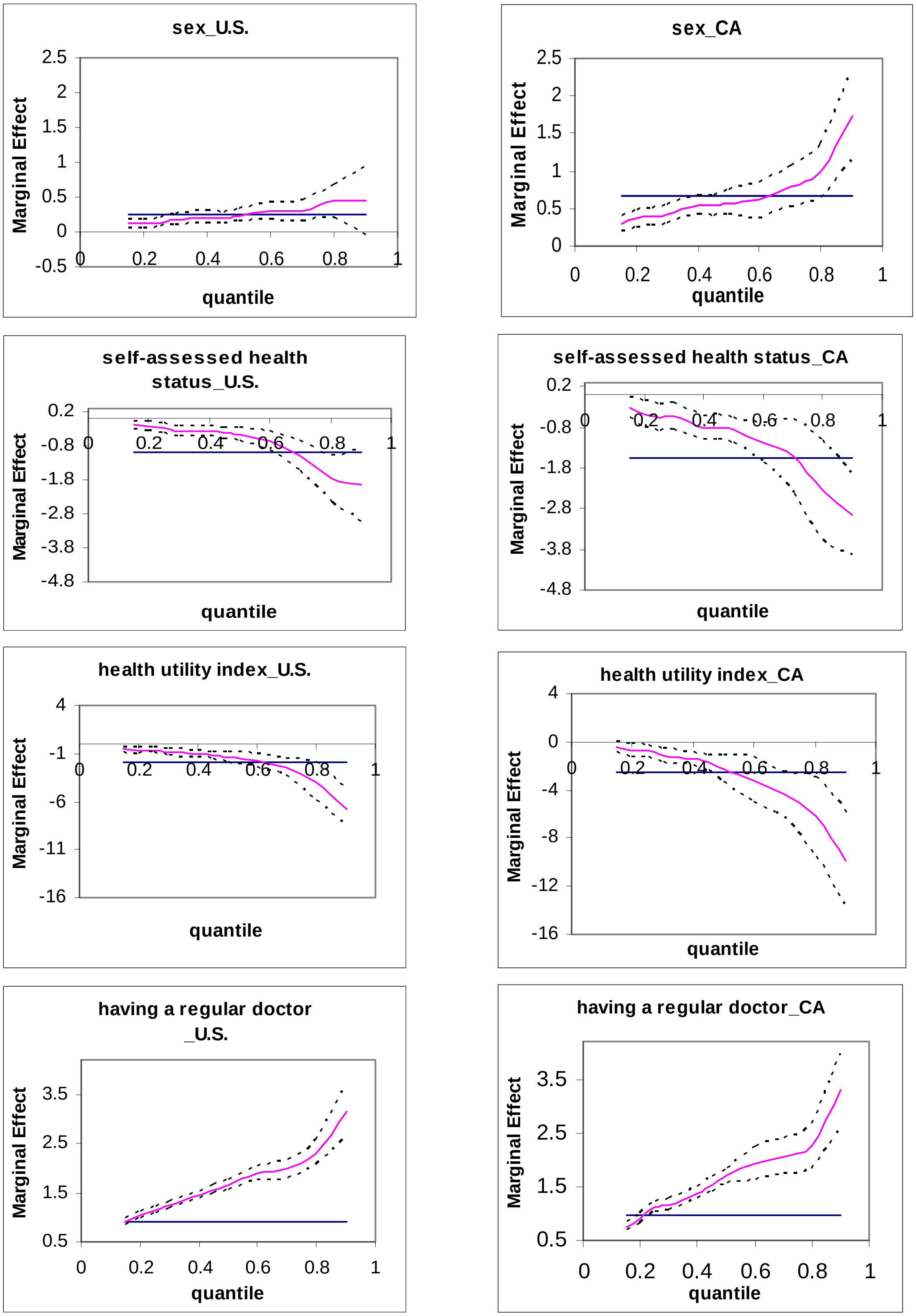
sex_U.S.
Marginal Effect
2.5
2
1.5
1
0.5
0
-0.5
0
0.2
0.4
0.6
0.8
1
quantile
sex_CA
Marginal Effect
2.5
2
1.5
1
0.5
0
0
0.2
0.4
0.6
0.8
1
quantile
self-assessed health status_U.S.
Marginal Effect
0.2
-0.8
-1.8
-2.8
-3.8
-4.8
0
0.2
0.4
0.6
0.8
1
quantile
self-assessed health status_CA
Marginal Effect
0.2
-0.8
-1.8
-2.8
-3.8
-4.8
0
0.2
0.4
0.6
0.8
1
quantile
health utility index_U.S.
Marginal Effect
4
-1
-6
-11
-16
0
0.2
0.4
0.6
0.8
1
quantile
health utility index_CA
Marginal Effect
4
0
-4
-8
-12
-16
0
0.2
0.4
0.6
0.8
1
quantile
having a regular doctor _U.S.
Marginal Effect
3.5
2.5
1.5
0.5
0
0.2
0.4
0.6
0.8
1
quantile
having a regular doctor_CA
Marginal Effect
3.5
2.5
1.5
0.5
0
0.2
0.4
0.6
0.8
1
quantile

Figure 22