THREE ESSAYS IN EMPIRICAL HEALTH ECONOMICS

By

Sisira Kumar Sarma

A Thesis Submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department of Economics
The University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES
*****
COPYRIGHT PERMISSION

THREE ESSAYS IN EMPIRICAL HEALTH ECONOMICS

BY

Sisira Kumar Sarma

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Doctor of Philosophy

Table of Contents

Abstract

It is well known that understanding the underlying process of the demand for health and health care utilization is crucial for a better assessment of the role of public intervention in the health sector. This thesis explores several theoretical and empirical aspects of these issues using unexplored and unique data sets, an appropriate economic theoretical framework, and advanced econometric techniques.

Essay 1 examines the factors determining the utilization of different types of health care from recent Canadian National Population Health Survey conducted by Statistics Canada. It uses the number of visits to GPs, specialists, and dentists and the number of nights spent in hospital as measures of utilization of health care. An intuitively appealing economic framework, in which individuals maximize the net benefits of visits, is used to base the analysis of health care utilization. The most striking results from this paper are that supplemental health insurance increases outpatient health visits, that there is vertical equity in the utilization of health care, and that there are some indications of supplier induced demand for health care. It is also found that *ex ante* and *ex post* utilization are two distinct stochastic processes.

Essay 2 examines the issues relating to demand for health care in rural India where much of the health services are typically provided at little or no *monetary* cost. This essay uses a discrete choice model to explain the underlying determinants of the demand for outpatient health care in rural India based on National Sample Survey (NSS) data for the first time. As opposed to fixed choice sets used in the literature, a variable choice set is constructed and used in this study to reflect the true choice generating process as close as possible. The relevant price data for unchosen alternatives in the choice set are imputed. The paper discusses econometric methods relating to identification, scaling, invariance, and consistency with the utility maximization

hypothesis that underlies the basis of modelling health care demand. Contrary to many earlier studies on the demand for health care in developing countries, prices and income are found to be statistically significant determinants of health care choice. Distance is a pronounced inhibiting factor in the demand for outpatient health care in rural India.

Linking Aging In Manitoba (AIM) longitudinal study on 1971 cohort's interview data with home care admission data, essay 3 explores the underlying determinants of elderly living arrangements. It is found that home care admission (*ex ante* home care utilization) reduces the demand for nursing home and increases the demand for independent living. Loss of a spouse affects independent living negatively and both cohabiting and nursing home residence positively. The effect of age on nursing home residence is positive and on independent living and cohabiting is negative. Educated people are more likely to live independently than cohabit or enter an institution. Similarly, those who are healthy and satisfied in life are more likely to live independently instead of cohabiting or entering nursing homes. Those who lived longer in the community are more likely to live independently or cohabit rather than enter an institution. Home ownership is positively associated with both independent living and cohabiting. The results are suggestive of possible income related inequity in institutionalization.

Acknowledgements

I am thankful to my thesis advisor Professor Wayne Simpson for constant discussion, encouragement, congeniality and invaluable guidance (theoretical and technical advices) throughout my studies at the University of Manitoba. I would like to thank my advisory committee for their advice and insightful comments. I am indebted to Professor Adeshir Sepehri for many discussions, useful comments and insightful suggestions. Professor Henry Rempel provided useful comments on the second essay and has reshaped my thinking in many ways ever since I took his course on Human Resources.

I owe special thanks to Aging In Manitoba Principal Investigator, Professor Betty Havens and Madelyn Hall, both of whom not only facilitated data access but also provided helpful insights and comments. Without their cooperation and discussion, my third essay would not have been possible. I thankfully acknowledge the creation of panel data files by Henry Dyck, senior programmer at Aging In Manitoba. The Ethics Board (University of Manitoba) and Health Information Privacy Committee (Manitoba Health) have kindly approved the proposal of the third essay.

I am grateful to the external examiner, Professor Livio Di Matteo, for his time and effort in going through this thesis. He has provided constructive comments and valuable feedbacks.

My special thanks goes to late Professor Costas Nicolaou, Professor Henry Rempel, Professor John Loxley, Professor Irwin Lipnowski, Professor Norman Cameron, Professor Elizabeth Troutt, Professor Laura Brown and Professor Raj Dhruvarajan for their mentorship. Gary Strike at the Dafoe Library has been very helpful in providing data sources available at the University of Manitoba. I am also thankful to many professors and fellow graduate students, especially Sharanjit Uppal and Rosmy Jean Louis at the Department of Economics for having influenced me to complete my thesis in many ways.

I am thankful to Professor William Greene for helpful comments while estimating discrete choice models using NLOGIT software package. Dr. Brian Poi of STATA Institute has been very helpful in providing many technical supports while using STATA software package.

I greatly appreciate the congeniality and superb service of the Department of Economics support staff, namely, Betty McGregor, Jean Wilson, Judy Ings, and Laura Ross. They are just wonderful people and willing to help in whatever capacity they could.

Canadian Commonwealth Scholarship from the International Council for Canadian Studies (ICCS), University of Manitoba top-up award and J. G. Fletcher award for thesis research are all gratefully acknowledged. Without the Commonwealth Scholarship, graduate studies at the University of Manitoba would have been an unthinkable proposition for me.

Numerous travel Awards from the University of Manitoba and Canadian Economics Association have helped to participate and present my papers in many national and international conferences. I would like to thank the participants of the 36th and 37th Annual Conferences of the Canadian Economics Association held at the University of Calgary and Carleton University, respectively. The comments of Professor Ian Irvine on my 2nd paper at the Canadian Economics Association Meetings are very helpful for publication.

Finally, my thanks go to my family. Nothing has stopped my mum and dad to pray for God for me to succeed in life. I am thankful to God for His blessings on my journey in life. My brother and sisters are my best friends and primary sources of encouragement, inspiration and strength in life.

The usual disclaimers apply.

1 Introduction

The Canadian health care system provides universal and comprehensive health coverage for all medically necessary hospital, in-patient, and outpatient physician services. The most remarkable features of the Canadian health care system are that it excludes third party private insurance for those services provided under the provincial health plans, it is publicly financed through taxation but provided largely by not-for-profit private agencies, and physicians are mostly paid for on a fee-for service. Although the management and delivery of health services are the responsibility of the respective provincial governments, the health system is often referred to as the national health insurance system because provincial health plans are linked to the principles of Medicare[1] at the federal level. However, pharmaceuticals outside hospitals are not part of the national health insurance plans.

The health status of Canadian Society has improved significantly over time. Life expectancy at birth has steadily increased from 69 years in the 1950s to 79 years in 1997. The incidence of low birth weight (less than 2,500 grams) has decreased to 5.8% during 1996 from 7.2% in 1961. The infant

[1]The mandate of Canada Health Act was to assure universality, comprehenesiveness, equitable access, public administration, and portability.

mortality rate (per 1,000 live births) has decreased dramatically from 27.3 in 1960 to 5.8 in 1996. The perinetal mortality rate[2] (per 1,000 total births) has steadily decreased from 28.4 to 6.7 between 1960 and 1996. Deaths due to communicable diseases are almost negligible. These impressive health outcomes are largely responsible for Canada being at the top of the human development index published by the UN for several years. However, recent instabilities within the health care system are posing a serious threat to the continuity and efficient functioning of the national health insurance scheme.

Under the Canadian national health insurance plan, a person who feels the need for medical care can visit the family physician or other health professionals of choice. The general practitioners are mostly the initial contact person for the patient, and in some sense are the controller of subsequent treatments - the demand for specialist services, hospital admissions, diagnostic testing, prescription drugs, etc. Since there are no substantial monetary costs at the point of access, it is interesting to look into the patients' behaviour in contacting a particular type of health professional, and the extent of its utilization. In practice, however, non-market forces like waiting time,

[2]Perinetal mortality is defined as the annual number of stillbirths and early neonatal deaths (deaths in the first week of life) per 1,000 total births (includes stillbirths). Stillbirths are defined here as gestational age of 28 or more weeks.

travel time, quality of services, and effectiveness of treatment might play an important role in choosing a health care professional. Contacting a general practitioner may be influenced by additional factors like the admitting privileges of the general practitioner and the relationship between a particular general practitioner and specialists (i.e., reputation for good referrals).

Under Canada's publicly financed and privately delivered health care system, patients pay little or none of the costs of their care.[3] Perhaps because of publicly funded health care and subsidies through tax exemptions for private supplemental health insurance, there has been increased demand for health services, thereby putting the Canadian health system under jeopardy and necessitating government regulations to contain costs. In order to regulate demand for health care, physician fees are controlled, specialist services are increasingly rationed and the introduction of new technology has been delayed substantially.[4] Canadians have less difficulty in seeing a general practi-

[3] This is true for contacting a physician, and to a limited extent for other health care professionals. However, purchase of prescription drugs (except inpatient prescription drugs), regular eye checkups, eye glasses or contact lenses, dental health care, private rooms during hospitalization days, etc. are basically out of pocket expenditures unless these are covered through the supplemental private health insurance plans. But private supplemental health insurance is typically available through the employer either at a zero cost or at a small cost due to both risk pooling and tax exemptions by the government. This has severe implications for both adverse selection and moral hazard in the utilization of health care. Stabile [43] examined the effects of tax exemptions to employer provided insurance on the utilization of health care in Canada, and found that additional health insurance policies lead to moral hazard in the use of health care.

[4] Compared to Canada, the US has 10 times as many magnetic resonance imaging (MRI)

tioner physician, but long waiting lists to see a specialist are well documented across the provinces.[5] As a result of the pressure on the health care system, there has been debate about reforming the Canadian health care system. In fact, health care system reform is not confined to Canada; nations around the world are facing similar dilemmas. Almost all the countries in the world are debating about who should pay for what and how best to organize and deliver health services so as to allocate scarce resources efficiently and work towards a healthier society. In this context, one way to attain a healthier society is to provide appropriate services at the right time regardless of individual ability to pay. Thus, a comprehensive study of the use and intensity of utilization of health services would assist the policy makers to address

units per capita; 3 times as many computerized axial tomagraphy (CAT) scanners; almost 3 times as many lithotripsy units; 3 times as many open heart surgery units; and 11 times as many cardiac catheterization units. (figures are collected from various newspaper and magazine sources in 1990's).

[5] Official statistical estimates show that 1,379,000 people in Canada are waiting for some medical service ranging from a visit to their general practitioner to hospital admission. Some of the annoying waiting statistics are as follows. The average waiting time to see an eye specialist in Prince Edward Island is around six months and another six months for treatment. It takes around seven weeks to see a gynecologist in New Brunswick and six months for treatment. To see an ENT (ear, nose, and throat) specialist, it takes around two weeks in Newfoundland, and another six months for treatment. In British Columbia, the waiting time for certain procedures like, cholecystectomies, prostatectomies, hip replacements, surgery for hemorrhoids and varicose veins, etc. are almost an year. In Ontario, the average waiting time for a CAT scan is around six months, MRI scan is around four months, eye surgery and orthopedic surgery is around one year. On an average, during 1993, it works out to be around five weeks to see a specialist in Canada. The waiting time for actual treatment is even longer (figures are collected from various newspaper and magazine sources during 1990's).

relevant public policy issues.

More specifically, we try to find answers to the following set of interrelated fundamental questions in this paper. What are the factors determining utilization of health care in Canada? Does supplemental insurance policy matter in health care utilization? Is there any indication of supplier-induced demand? How does health care utilization responds to the need? Does income matter in health care utilization?

The paper is organized as follows. We briefly review the literature concerned with the demand for health care as the basis for health care utilization in section 2. We develop an intuitive theoretical framework for studying the utilization of health care in section 3. Section 4 reviews the econometric methods and test procedures for the empirical exercise. We explain our data source and variable information in section 5. In section 6, we report and discuss our results. Finally, section 7 discusses our conclusions and the limitations of this study.

2 Literature Review

Many theoretical and empirical studies of the demand for health care consider the patient as the sole agent (Grossman [19], [20], Muurinen [39], Wagstaff [47], etc.). However, recent literature suggests a need to separate the modelling of patient-initiated contact from the intensity of use of health care (Zweifel [52], Manning et al. [34], Pohlmeier and Ulrich [41], and Gerdtham [14]). There are essentially two stages involved here. In the first stage, the patient presumably initiates a contact decision whenever she decides to visit her family physician or any other health professional. In the second stage, it may be either the health professional or both the patient and health professional together[6] who decide the intensity of use of health services. The intensity of the use of health services could be in terms of subsequent visits, diagnostic procedures, further treatment, referral to a specialist, recommendation for hospitalization, surgical procedures, etc.

Once a person decides to visit a health professional, however, it remains to determine the extent of utilization. The Grossman model ([19], [20]) implies that the individual is the prime decision maker regarding the use of health

[6]It is conceivable that the patient might influence the physician's decision about the length of hospitalization stay, or might seek additional treatment or subsequent visits.

care. Within the count data tradition, the empirical studies of Cameron et al. [4], Dev and Trivedi [10] and Dev and Trivedi [11] are consistent with Grossman's model.[7] Zweifel [52] on the other hand, advocates a principal-agent framework in which the patient may initiate the contact, but physicians have much authority in deciding the subsequent treatment. Similar ideas are also reflected in the work of Manning et al. [34]. Pohlmeier and Ulrich [41] and Gerdtham [14] actually implemented the principal-agent framework empirically, including the statistical refinements involving count data techniques developed by Mullahy [36].

However, the Grossman model and the agency model might offer a com-

[7] It is to be noted that some predictions of the Grossman model are contradicted by empirical evidence. Wagstaff's [47] study of the demand for health in the Danish Welfare Survey is a starting point for criticism. Wagstaff used multiple-indicators-multiple-causes (MIMIC) techniques to estimate a multidimensional version of the structural and reduced form of the demand for health equation. However, the structural equation that he derives at the end is known as the conditional output demand function (see Grossman [21]). The latent variable health in his MIMIC formulation is a positive correlate of good health, and good health is one of the regressors in the conditional demand function for physician visits. The estimated results from the reduced form equation of his pure investment model are consistent with theoretical predictions. Similar cross-sectional results are also found in Grossman [20], Muurinen [39], Gredtham et al. [15] , and Gredtham and Johanesson [15]. However, the estimated parameters of the structural model are of the wrong sign. More importantly, the coefficient of good health in the conditional demand equation for physician visits is negative, a contradiction to the theory. Erbsland et al. [13] also found similar evidence. In fact, these findings have been the basis for the criticism of Grossman's model by Zweifel and Breyer [53] who conclude: *"Many of the implications of the Grossman model are contradicted by available empirical evidence. Most important the notion that expenditures on medical care constitutes a demand derived from an underlying demand for health cannot be upheld because health status and demand for medical care are negatively rather than positively related"*. Grossman's model has also been attacked by Zwcifel and Breyer [53] on the basis of incorrect signs of coefficients in many studies.

plementary explanation for the utilization of health care. A Grossman style interpretation might be appropriate for explaining the contact decision, whilst an agency approach might be suitable for the interpretation of the frequency decision. In fact, this approach is supported by evidence in Pohlmeier and Ulrich [41]. In their two-part model, they found that physician density does not affect the contact decision while it affects the frequency decision positively. Thus, there is some justification for supplier-induced demand in the utilization of health care. As discussed earlier, a common feature of the Canadian health insurance system is that individuals pay little at the point of access, and data on out-of-pocket spending is simply not available or unimportant due to extensive use of supplemental health insurance in Canada. Therefore, the monetary costs can be typically captured through supplemental private health insurance status (see Pohlmeier and Ulrich [41] and Dev and Trivedi [10]), or through the coinsurance rate in Dev and Trivedi [11]. The effect supplemental private health insurance is positive on the contact decision but not on the frequency decision in Pohlmeier and Ulrich [41]. Similarly, higher copayment rates result in lower probability of contact, but not frequency of visits, in Dev and Trivedi [11]. This effect of health insurance closely resembles the notion of *ex ante* moral hazard by Zweifel and Manning [54]. In the next section, we develop an intuitive framework to model contact and frequency decisions.

3 The Framework

Let us assume that an individual, denoted as i visits the jth health professional $h_{ij} \in H_{ij}$ times in a given period of time to meet her physical, emotional or mental health needs. Assume that H_{ij} is a non-empty, compact subset of a finite dimensional euclidean space, $H_{ij} \in [0, \Re_{+}) \forall i, j$. Further, assume that the individual has a von Neumann-Morgenstern utility function $U(h_{ij}, \mathbf{X}_i, S_h)\ \forall i, j$, which depends on the number of visits h_{ij}, a set of socioeconomic and health backgrounds $\mathbf{X}_i$, and the state of health S_h. Assume that there could be many possible states of the world and only one of these, say $s_{h_i} \in S_h$ is realized for individual i in a given time period. Complete knowledge of the realized state of health $s_{h_i} \in S_h$ may be completely known to both the patient and health care professional or partially known to the patient but completely known only to health professional. In either case, the patient must take a decision about whether to visit a health care professional and, if so, how many times and how much to consume other medical care resources. However, contacting any health care professional does involve significant costs, both monetary (like the purchase of drugs) and non-monetary (travel, time, and other opportunity costs). Similarly, there are substantial benefits from the consumption of health care, including the augmented human capital and effective labour in a given state of morbidity. Different patients might possibly have a different number of visits to different health care professionals when confronted with the same state of health, because the subjective valuation that they place on both the costs and the benefits may differ. The differences in the valuation depend on the specific characteristics of the patient, only some of which are observed by the researcher. Formally, let the cost of utilization of h_{ij} visits to the jth type of health care by the ith person in state $s_{h_i} \in S_h$ be $C(h_{ij}, \mathbf{X}_i, \varepsilon_i)$,where ε_i is the random element associated with the cost function. Similarly, the benefit function is $B(h_{ij}, \mathbf{X}_i, \boldsymbol{\eta}_i)$, where $\boldsymbol{\eta}_i$ is the random element associated with the benefit function. Assume that the utility function is additively separable in benefits and costs. Thus, we have

$$U(h_{ij}, \mathbf{X}_i, S_h) = B(h_{ij}, \mathbf{X}_i, \eta_i) - C(h_{ij}, \mathbf{X}_i, \varepsilon_i), \forall i, j. \tag{1}$$

Assumption 1: $\Delta_h B(h_{ij}, \mathbf{X}_i, \boldsymbol{\eta}_i) > 0, \Delta_h^2 B(h_{ij}, \mathbf{X}_i, \boldsymbol{\eta}_i) < 0 \forall i, j$. That is, the benefit function is increasing at a decreasing rate in the number of visits for every individual, and for all types of health services.

Assumption 2: $\Delta_h C(h_{ij}, \mathbf{X}_i, \varepsilon_i) > 0, \Delta_h^2 C(h_{ij}, \mathbf{X}_i, \varepsilon_i) > 0 \forall i, j$. That is, the cost function is increasing at an increasing rate in the number of visits for every individual, and for all types of health services.

Thus, the patient's net benefit is actually not completely determined by $h_{ij} \backslash \mathbf{X}_i$; rather it is stochastically affected by the state of health, $s_{h_i} \in S_h$. Let, $P_s(h_{ij})$ be the probability that the state $s_{h_i} \in S_h$ has occurred resulting in h_{ij} visits to the jth health professional by the ith individual.

Assumption 3: $P_s(h_{ij}) \geq 0 \forall s_{h_i} \in S_h, h_{ij} \in H_{ij} \forall i, j$.

Let $F_s(h_{ij}) = \sum_{k=1}^{S} P_k(h_{ij})$ be the corresponding distribution function. The expected utility of individual i from the consumption of the jth type of health care from h_{ij} visits can therefore be written as

$$U(h_{ij}, \mathbf{X}_i, S_h) = \sum_{k=1}^{S} P_k(h_{ij}) \left[B(h_{ij}, \mathbf{X}_i, \boldsymbol{\eta}_i) - C(h_{ij}, \mathbf{X}_i, \varepsilon_i) \right]. \qquad (2)$$

The optimal number of visits to the jth health care professional by the ith individual is thus the solution to the following maximization problem:

$$\max_{\{h_i j\}} \sum_{k=1}^{S} P_k \left(h_{ij}\right) \left[B \left(h_{ij}, \mathbf{X}_i, \boldsymbol{\eta}_i\right) - C \left(h_{ij}, \mathbf{X}_i, \varepsilon_i\right)\right] \forall, j. \tag{3}$$

The solution to (3) must be at a point where the expected marginal benefit is equal to expected marginal cost. Assume that $B(.)$ and $C(.)$ are multiplicatively separable in each of their arguments. That is, $B\left(h_{ij}, \mathbf{X}_i, \boldsymbol{\eta}_i\right) \equiv B\left(h_{ij}\right) \phi\left(\mathbf{X}_i\right) \boldsymbol{\eta}_i$, and $C\left(h_{ij}, \mathbf{X}_i, \varepsilon_i\right) = C\left(h_{ij}\right) \psi\left(\mathbf{X}_i\right) \varepsilon_i$. Suppose that a state $s_{h_i} \in S_h$ has occurred. The individual then decides what type of health care to consume and how much so as to maximize the expected net benefit. For a simple analysis, assume that $h^*_{ij} > 0$ is the equilibrium number of visits to the jth health professional by the ith person when the state $s_{h_i} \in S_h$ is known. This means that $\left(h^*_{ij} - 1\right)$ might not be optimal for the ith person in that state. This would imply that

$$\left\{ \begin{array}{c} B\left(h^*_{ij}\right) \phi\left(\mathbf{X}_i\right) \boldsymbol{\eta}_i - B\left(h^*_{ij} - 1\right) \phi\left(\mathbf{X}_i\right) \boldsymbol{\eta}_i \geq 0 \text{ if } h^*_{ij} > 0 \\ B\left(h^*_{ij}\right) \phi\left(\mathbf{X}_i\right) \boldsymbol{\eta}_i - C\left(h^*_{ij}\right) \psi\left(\mathbf{X}_i\right) \varepsilon_i \leq 0 \text{ if } h^*_{ij} = 0 \end{array} \right\}. \tag{4}$$

In order to analyze this problem, define the net benefit of visiting the jth health professional h_{ij} by the ith person as a latent variable HF^*_{ij} such that

$HF_{ij}^{*} = \Phi\left(\mathbf{X}_i\right) \mathrm{e}_i$, where, e_i captures the unobserved heterogeneity that is unknown to the researcher (i.e., e_i captures unobserved heterogeneity arising from both $\boldsymbol{\eta}_i$ and ε_i). Define an indicator function HF_{ij}^{*} such that

$$\left\{ \begin{array}{c} h_{ij} = 1 \text{ if } HF_{ij}^{*} > 0 \\ \\ h_{ij} = 0 \text{ if } HF_{ij}^{*} \leq 0 \end{array} \right\}. \tag{5}$$

Equation (5) says that if individual i expects a strictly positive net benefit from visiting the jth health professional $\left(HF_{ij}^{*} > 0\right)$, then she will consult h_{ij} times, otherwise $HF_{ij}^{*} \leq 0$ and she will not consult $\left(h_{ij} = 0\right)$. The necessary conditions for optimality are:

$$\left\{ \begin{array}{c} \dfrac{B\left(h_{ij}^{*}\right) - B\left(h_{ij}^{*} - 1\right)}{C\left(h_{ij}^{*}\right) - C\left(h_{ij}^{*} - 1\right)} = \frac{\psi(\mathbf{X}_i)\varepsilon_i}{\phi(\mathbf{X}_i)\eta_i} > 0 \text{ if } HP_{ij}^{*} > 0 \\ \dfrac{B\left(h_{ij}^{*}\right)}{C\left(h_{ij}^{*}\right)} = \frac{\psi(\mathbf{X}_i)\varepsilon_i}{\phi(\mathbf{X}_i)\eta_i} = 0 \text{ if } HP_{ij}^{*} \leq 0 \end{array} \right\} \tag{6}$$

Different discrete probability models may be employed to model equation (6). However, in reality the term $\left(\frac{\psi(\mathbf{X}_i)\varepsilon_i}{\phi(\mathbf{X}_i)\eta_i}\right)$ could be negative. In order to make sure that this term is non-negative, we need some transformation. Clearly the exponential transformation of the form $\exp\left(\mathbf{X}_i\boldsymbol{\beta}\right)$ serves this purpose, where $\boldsymbol{\beta}$ is to be estimated. However, the important question that needs to be addressed is the underlying decision making processes in determining h^*_{ij}.

First, it might be possible that the patient is fully informed about the realization of $s_{h_i} \in S_h$ and knows where to go and what to purchase. Thus, the individual is the sole decision maker about whether to contact a health professional and how often, reflected in the number of follow-up visits (basically, the Grossman interpretation).

Second, it is conceivable that the patient may not be fully informed about the realization of $s_{h_i} \in S_h$ when she initiates a contact to buy information and subsequent treatment. A health professional might act solely in the best interest of patient and provide adequate information according to the prevailing technology and principles governing professional medical ethics. It is also quite possible that a health professional could exploit the asymmetric information and take informational advantage by inducing more utilization than would have been the situation if the patient had the same knowledge. In the second stage, it is mostly up to the health professional to determine the actions to be taken and the number of visits so as to maximize her own interests rather than that of the patient. This approach stems from the conventional principal-agent framework, where the agent maximizes her own interests and presumably may not act in the best possible interest of her principal. This is referred to as the hurdle model in the literature (Mullahy [36]). Since, once a patient decides to contact a health professional, the hurdle is crossed. The hurdle model is also known as the two-part model in the literature. Thus the essence of this is that the final determination of h_{ij}^* essentially involves two separate stages. The first stage occurs when the patient decides to contact a health professional (the transition stage, i.e., the transition from 'no contact' to 'contact' or from 'zero-state' to some 'non-zero-state'). The second stage can be characterized by the intensity of utilization. In terms of the statistical refinements, there can be excess zeros, that is, there are many more zeros than is consistent with the count regression models.

4 Econometric Methods

In this section, we discuss the underlying data generating process for non-negative integer outcomes and the possible econometric techniques that can be employed to model health care utilization. Since the outcome of interest is necessarily a non-negative integer, with many zeros in some instances, a discrete probability distribution provides a natural theoretical basis for analysis rather than the conventional normal distribution underlying OLS regression. The Poisson regression model is thus a quite natural starting point for our analysis. In the context of our analysis, we are interested in modelling the probability of making h_{ij} visits to the jth health professional by the ith individual in the state $s_{h_i} \in S_h$ during a given time interval. As assumed earlier, there are a large number of possible health states S_h, and only one state $s_{h_i} \in S_h$ is realized leading to h_{ij}^* visits.

4.1 Poisson and Negative Binomial Models

Count data techniques are discussed in detail in Hausman, Hall, Griliches [25], Cameron and Trivedi [?] [6], Cameron, Trivedi, Milne, and Piggot [?], and Gurmu and Trivedi [22]. If the random variable h_{ij} is Poisson distributed with intensity parameter μ_{ij} during the time length t then h_{ij} has the density,

$$\Pr\left(h_{ij} = h_{ij}^*\right) = \frac{e^{-\mu_{ij}t}\left(\mu_{ij}t\right)^{h_{ij}^*}}{h_{ij}^*!}, h_{ij}^* = 0, 1, 2, ... \tag{7}$$

If we set the length $t = 1$, then the Poisson density is given by

$$\Pr\left(h_{ij} = h_{ij}^*\right) = \frac{e^{-\mu_{ij}}\mu^{h_{ij}^*}}{h_{ij}^*!}, h_{ij}^* = 0, 1, 2, ... \tag{8}$$

where, μ_{ij} is the parameter to be estimated. When there are exogenous variables, for the reasons mentioned in section (3), μ_{ij} is modelled as $\mu_{ij} = \exp(\mathbf{X}_i\boldsymbol{\beta})$ where $\mathbf{X}_i$ is the vector of explanatory variables and $\boldsymbol{\beta}$ is the vector of parameters to be estimated. If the data generating process for h_{ij} indeed follows a Poisson distribution with mean μ_{ij} then maximum likelihood estimation theory implies that, $\widehat{\boldsymbol{\beta}}_p \overset{a}{\sim} N\left(\boldsymbol{\beta}, \mathbf{V}_{ml}\left(\widehat{\boldsymbol{\beta}}_p\right)\right)$.

Although the Poisson model defined in (8) is attractive to model non-negative integer outcomes, it has a couple of weaknesses. The first is the equidispersion property; that is $E(h_{ij} \mid \mathbf{X}_i, s_{h_i} \in S_h) = Var(h_{ij} \mid \mathbf{X}_i, s_{h_i} \in S_h) = \mu_{ij}$. In empirical work, this property often does not hold, and the Poisson model fails to account for overdispersion or underdispersion in the data. Imposition of this restriction usually yields consistent estimates of the mean parameters but the effect on standard errors and t-statistics could be substantial, generally yielding small estimated standard errors of $\widehat{\beta}$ (Cameron and Trivedi [?], and predicting the number of zeros incorrectly. Note that in the raw data, the unconditional mean $(E(h_{ij}))$ is strictly greater than the unconditional variance $(Var(h_{ij}))$ for all types of health care utilization considered in this study[8] implying that the raw data do exhibit overdisper-

[8]See Tables 2a and 2b.

sion. However, this does not necessarily rule out the use of Poisson regression unless $Var(h_{ij} \mid \mathbf{X}_i, s_{h_i} \in S_h) > E(h_{ij} \mid \mathbf{X}_i, s_{h_i} \in S_h)$. So, there is clearly a need for test procedures and alternative modelling techniques to account for overdispersion.

Second, the Poisson model assumes that events occur independently over time. However, in real life there might be some form of dependence between successive events. The independence assumption in our context implies that the probability of the nth visit to the jth health care professional is independent of the $(n+1)$th and $(n-1)$th visits. This is clearly a very restrictive assumption and even inconsistent with the commonly observed facts about dynamic dependence.

In order to test the null hypothesis of equidispersion, the likelihood ratio test proposed by Cameron and Trivedi [5] is used. The null hypothesis is expressed as $H_0 : Var(h_{ij} \mid \mathbf{X}_i, s_{h_i} \in S_h) = \mu_{ij}$ and compared with two alternative hypotheses (1) $H_1 : \mu_{ij} + \alpha\mu_{ij}$, and (2) $H_1 : \mu_{ij} + \alpha\mu_{ij}^2$. The test for overdispersion is thus a test for $\alpha = 0$. If the null hypothesis is rejected, which means that the dependent variable displays overdispersion, a more flexible discrete probability distribution is clearly needed, such as the Negative Binomial distribution (Johnson and Kotz [28], Hausman, Hall, and

Griliches [25], and Cameron and Trivedi [3]).

In situations where data exhibits overdispersion and the independence assumption is expected to be violated the negative binomial model provides close approximation to the underlying true model (Johnson and Kotz [28], Cameron and Trivedi [6]), provided that there are no other complications like selection bias, endogenous regressors, etc. The Negative Binomial regression model also allows for cross-sectional heterogeneity by introducing an unobserved individual effect into the conditional mean function. That is, $\mu_{ij} = \exp(\mathbf{X}_i\boldsymbol{\beta})\mathbf{e}_i$, where $\mathbf{e}_i$ could possibly represent either cross-sectional heterogeneity as evident in most micro data or specification errors (see for instance, Heckman [27]). If $\mathbf{e}_i$ is a random variable with expected value one and variance α characterizing the unobserved cross-sectional heterogeneity, then $E\left(h_{ij} \mid \mathbf{X}_i, \mathbf{e}_i, s_{h_i} \in S_h\right) = \mu_{ij}$ and $Var\left(h_{ij} \mid \mathbf{X}_i, \mathbf{e}_i, s_{h_i} \in S_h\right) = \mu_{ij} + \alpha\mu_{ij}^2$.

The negative binomial (NB) density function is given by:

$$\Pr\left(h_{ij} = h_{ij}^*\right) = \frac{\Gamma\left(h_{ij}^* + \nu\right)}{\Gamma\left(h_{ij}^* + 1\right)\Gamma\left(\nu\right)}\left(\frac{\nu}{\nu + \mu_{ij}}\right)^{\nu}\left(\frac{\mu_{ij}}{\nu + \mu_{ij}}\right)^{h_{ij}^*}, h_{ij}^* = 0, 1, 2, ... \tag{9}$$

where $\Gamma\left(.\right)$ is the gamma distribution function, $\mu_{ij} = \exp\left(\mathbf{X}_i\boldsymbol{\beta}\right)$, $\nu = \frac{\mu_{ij}}{\alpha} = \alpha^{-1}\left(\exp\left(\mathbf{X}_i\boldsymbol{\beta}\right)\right)$, or $\nu = \frac{1}{\alpha}$, and $\alpha \geq 0$. In equation (9), if $\alpha = 0$, the model reduces to a Poisson model. If $\nu = \frac{\mu_{ij}}{\alpha}$, the model is referred to as Negative Binomial-I (NB1) model, if $\nu = \frac{1}{\alpha}$, it is known as Negative Binomial-II (NB2) model.[9] For the NB1 density, $Var\left(h_{ij} \mid \mathbf{X}_i, \mathrm{e}_i, s_{h_i} \in S_h\right) = \mu_{ij} + \alpha\mu_{ij}$. In the NB2 density, $Var\left(h_{ij} \mid \mathbf{X}_i, \mathbf{e}_i, s_{h_i} \in S_h\right) = \mu_{ij} + \alpha\mu_{ij}^2$. As indicated earlier in this section, if the null hypothesis $H_0 : \alpha = 0$ is rejected, one can choose either a NB1 or a NB2 model, depending on the log likelihood.

4.2 Hurdle Models

Hurdle models are increasingly used in health economics literature (Mullahy [36], Pohlmeier and Ulrich [41], Gurmu and Trivedi [22]). The hurdle model is interpreted as a two part model. The first part models the probability that the threshold is crossed, the second part is a truncated count data model. The idea behind the hurdle model is that a binomial probability model governs the realization of a zero or a non-zero outcome. As before, the first stage can be interpreted as the contact decision by the patient. The data generating process for the second stage might be significantly different from the first stage. A logit or probit model is usually employed for the first part, and a truncated Poisson or a truncated Negative Binomial model

[9]In this paper, we often use then term NB for NB2 model. Whenever Negative Binomial-1 is used, we use the term NB1.

is usually employed for the second part of the two part model.[10] Although the zero-truncated Poisson or zero-truncated Negative Binomial model is employed in most applications, the threshold need not necessarily be at zero.

From the Poisson formulation, we know that $\Pr(h_{ij} = 0) = \exp(-\mu_{ij})$ and $\Pr(h_{ij} > 0) = 1 - \exp(-\mu_{ij})$. Since the probability of a zero in a zero-truncated process is zero, we are interested to know $\Pr(h_{ij} = h_{ij}^* \mid h_{ij} > 0)$. By using the rules of conditional probability ($\Pr(A \mid B) = \Pr(A \cap B) / \Pr(B)$), the zero-truncated density function is given by:

$$\left\{ \begin{array}{c} \Pr\left(h_{ij} = h_{ij}^* \mid h_{ij} > 0, \mathbf{X}_i, \mathbf{e}_i, s_{h_i} \in S_h\right) = \frac{\Pr\left(h_{ij} = h_{ij}^* \mid \mathbf{X}_i, \mathbf{e}_i, s_{h_i} \in S_h\right)}{\Pr(h_{ij} > 0)} = \\ \frac{\exp(-\mu_{ij})\mu_{ij}^{h_{ij}}}{\left[1 - \exp(-\mu_{ij})\right] h_{ij}!} \end{array} \right\}. \tag{10}$$

Let π and $1 - \pi$ be the probability of clearing and not clearing the hurdle respectively. Then

$$\Pr\left(h_{ij} = h_{ij}^* \mid \mathbf{X}_i, \mathbf{e}_i, s_{h_i} \in S_h\right) = \left\{ \begin{array}{c} 1 - \pi \text{ if } h_{ij} = 0 \\ \pi \frac{\exp(-\mu_{ij})\mu_{ij}^{h_j}}{\left[1 - \exp(-\mu_{ij})\right] h_{ij}!} \text{ if } h_{ij} > 0 \end{array} \right\}. \tag{11}$$

The zero-truncated Negative Binomial density function (Gurmu and Trivedi

[10]The Sample Selection Model is another competing econometric technique in this context. However, the hurdle model is attractive for a variety of reasons. An extensive debate over the two part model and the sample selection model can be found in Jones [29].

[22]) is specified as:

$$\left\{ \begin{array}{c} \Pr\left(h_{ij} = h_{ij}^{*} \mid h_{ij} > 0, \mathbf{X}_i, \mathbf{e}_i, s_{h_i} \in S_h\right) = \\ \frac{\Gamma\left(h_{ij}^{*} + \alpha^{-1}\right)}{\Gamma\left(h_{ij}^{*} + 1\right)\Gamma\left(\alpha^{-1}\right)} \left[\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_{ij}}\right)^{-\alpha^{-1}} - 1\right]^{-1} \left(\frac{\mu_{ij}}{\alpha^{-1}+\mu_{ij}}\right)^{h_{ij}^{*}}, \alpha \geq 0, h_{ij}^{*} = 1, 2, ... \end{array} \right\}. \tag{12}$$

In (12), if $\alpha = 0$, the model reduces to a zero-truncated Poisson model. Since the zero-truncated Poisson model is nested in the zero-truncated negative binomial model, a likelihood ratio test is applicable.

4.3 Zero Inflated Models

Suppose the event of interest moves from 'zero-state' to some 'non-zero-state'. The second stage is an event-count process such that the zero inflated Poisson (ZIP) or zero inflated Negative Binomial (ZINB) models are appropriate (Lambart [30]). The ZIP model allows for excess zeros and ZINB allows for both excess zeros and between-subject heterogeneity. Econometrically, this represents overdispersion through an excess of zeros; that is, there are many more zeros than is consistent with the count regression models, such as Poisson or Negative Binomial models. Since zeros have special economic significance and cannot be ignored econometrically, methods to deal with the structural zeros need to be considered.

Define a binary variable c to indicate zero and non-zero outcomes. We observe the underlying health utilization variable h_{ij} if $c = 1$; otherwise $c = 0$. Thus,

$$h_{ij} = \begin{Bmatrix} h_{ij} \text{ if } c = 1 \\ 0 \text{ if } c = 0 \end{Bmatrix}. \tag{13}$$

Let $\Pr(c = 1)$ be denoted as p and $\Pr(c = 0)$ is denoted as $1 - p$. In this framework, the probability of obtaining a zero outcome is:

$$\Pr(h_{ij} = 0) = \Pr(c = 0) + \Pr(c = 1, h_{ij} = 0) = (1 - p) + p \Pr(h_{ij} = 0). \tag{14}$$

The probability of obtaining non-zero outcome is:

$$\Pr\left(h_{ij} = h_{ij}^{*}\right) = \Pr(c = 1) \Pr\left(h_{ij} = h_{ij}^{*}\right) = p \Pr\left(h_{ij} = h_{ij}^{*}\right), \forall h_{ij} > 0. \tag{15}$$

The probability function for h_{ij} is:

$$\Pr(h_{ij}) = (1 - p)^{1-c} + p\left(f\left(h_{ij}\right)\right), \tag{16}$$

where $f(.)$ is some specified density function. If c is specified as a logit or probit model and $f(h_{ij})$ is specified as a Poisson density function, then we have the ZIP model. Alternatively, if $f(h_{ij})$ is specified as a Negative Binomial density then we have the ZINB model. In the original paper, Lambert [30] uses the logit model for c. For the ZINB specification, the second stage is characterized by the following probability density function:

$$\left\{ \begin{array}{c} \Pr\left(h_{ij} = h_{ij}^* \mid \mathbf{X}_i, \mathbf{e}_i, p, s_{h_i} \in S_h\right) = \\ \dfrac{\Gamma\left(h_{ij}^* + \alpha^{-1}\right)}{\Gamma\left(h_{ij}^* + 1\right)\Gamma\left(\alpha^{-1}\right)} \left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_{ij}}\right)^{\alpha^{-1}} \left(\frac{\mu_{ij}}{\alpha^{-1}+\mu_{ij}}\right)^{h_{ij}^*}, \alpha \geq 0, h_{ij}^* = 0, 1, 2, ... \end{array} \right\} \tag{17}$$

The ZINB model reduces to the ZIP model when $\alpha = 0$. Since ZIP is nested in ZINB, the likelihood ratio test is applicable. If the null hypothesis $H_0 : \alpha = 0$ is accepted, then there is a suspicion that the introduction of a separate 'zero-state' data generating process sufficiently accounts for overdispersion in the data, and there is no reason to allow for cross-sectional heterogeneity. On the other hand, cross-sectional heterogeneity is almost inevitable in micro data, which leads to some objections about whether it is at all necessary to model zeros separately. In order to answer this question, we need to test ZINB against NB models. Since, ZINB and NB models are not nested, a likelihood ratio test is not applicable. Because the two underlying distributions are specified, we can apply a test developed by Vuong [46]. The Vuong test statistic has an asymptotic standard normal distribution with large positive values favouring the zero inflated model (i.e., ZIP or ZINB) and large negative values favouring the non-zero inflated models (i.e., Poisson or NB).

4.4 Latent Class Models

The recent empirical evidence suggests that the hurdle model cannot separately identify the parameters governing two decision processes (Santos Silva and Windmeijer [42]). Another potential problem in the health care utilization data is that it is almost impossible to distinguish different illness spells during the one year period. Although the zero inflated models capture excess zeros, they only allow mixing with respect to zeros. However, the nature of an illness spell may affect both zero and positive outcomes. One way to capture this phenomena is to use latent class models (Dev and Trivedi [11]). Latent class models are based on the standard count data models (Poisson or NB) but they allow for modelling unobserved heterogeneity across individuals, splitting the population into different health groups. The intuition behind the latent class models framework in our context implies that the state of health s_h is unobservable to the researcher. The widely used proxy variables such as self assessed health status, chronic conditions, and disability days may not capture an individual's long-term health status. So, in the context of two latent classes, one may distinguish health care utilization between healthy and less healthy populations. The distribution function of the unobservable heterogeneity is approximated by using a finite mixture distribution (Heckman and Singer, [26]). In the formulation of Heckman and Singer [26], only the constant term varies across classes. However in this formulation, each class has its own parameter vector. The negative binomial model allows for a separate dispersion parameter in each class.

5 Data and Variable Specifications

The empirical work uses the data from National Population and Health Survey (NPHS) of 1998-99, conducted by Statistics Canada. The survey contains a wealth of socioeconomic, health profile, morbidity profile, and health care utilization information. Since a number of variables are not applicable to younger age cohorts, only individuals aged 15 or over are taken into account. Realistically speaking, most young persons (less than 15 years of age) do not make their own decisions to consult a health professional. Our focus is whether or not patients seek to consult a health professional and how often, i.e., the intensity of use. Since our analysis is confined to publicly available micro data files, household responses from the health file are extracted to study utilization of health care.

A sample of $17,244$ observations are available in the health micro data files of the NPHS 1998-99 survey. Most Canadians consulted a health professional at least once in the 12–month period, 93.231% during 1998-99, 76.044% during 1996-97, and 73.634% during 1994-95. Of those who consulted health professionals, physicians were the most dominant choices. After deleting the "not applicable" and "not reported" responses, the sample size is reduced to $9,793$.

The dependent variables in our analysis is the number of consultations to a health professional, which has been indexed by j. Each respondent was asked about how many times she consulted a health professional in the previous 12 months. In this paper, we considered five dependent variables – (a) **Doctor** – the number of visits to a doctor (general practitioner or specialist) during the year preceding the survey date; (b) **GP** – the number of visits to a general practitioner during the year preceding the survey date; (c) **Specialist** – the number of visits to a specialist during the year preceding the survey date; (d) **Dentist** - the number of visits to a dentist or orthodentist during the year preceding the survey date; and (e) **Nights** - the number of nights spent as patient during the year preceding the survey date. Table 1 presents the list of dependent and explanatory variables used in this paper. The frequency distribution and the moments of the raw data on all health care utilization variables are presented in Tables 2a and 2b, respectively. It is clearly evident from Table 2b that the raw data for all health care utilization variables are highly skewed. The explanatory variables include demographic, household, socioeconomic (i.e., predisposing), enabling, need, and life-style variables.

Demographic, Household, and Socioeconomic Variables: Demographic, household, and socioeconomic variables generally capture the indirect measures of morbidity for individuals of different age, sex, and socioeconomic backgrounds, etc. Gender is represented by a $0-1$ dummy (female =1, male = 0). Marital status is characterized by two dummy variables (currently married =1, zero otherwise) and (widows, separated, and divorced=1, zero otherwise), which implies that singles are the reference group. Three age dummies are included: 35 - 59 years, 60 - 74 years and 75 or older. So, the reference category for age is 15 - 34 years. Immigration status of the respondent is represented by three dummies: years of immigration less than 4 years, 5 - 9 years, and 10 years or more. Thus, the reference category for immigration status is those who are native born Canadians.

Enabling Variables: Enabling variables are basically the access indicators that facilitate utilization of a particular health care. Some of the enabling variables are representative of the indicators of supply of health care (like geographical location and different waiting times in different provinces) while others are representative of demand side indicators (income, education, and supplemental health insurance). The geographical location of the respondent is represented by two dummies: urban and metropolitan; the reference category is rural. The provincial heterogeneity about waiting times, structure, delivery, organizational set up, etc. is represented by a series of provincial dummies with Ontario as the reference category. Educational status of the respondent is characterized by two dummies: respondent completed the secondary education and respondent completed post secondary education; the reference category for educational status of the respondent is thus those who have less than secondary level education. Three income dummies are included: household income in the middle income quartile, household income in the upper middle income quartile, and the household income in the highest income quartile; the reference category for income is those individuals who belong to the lowest income quartile. Availability of supplemental insurance, either through employee based insurance or personal insurance, is characterized by three $0-1$ dummies: prescription, dental and hospitalization insurance.

Need Variables: Need for health care has been interpreted in several alternative ways, such as need for a stock of health, the capacity to benefit from medical care consumption, and the expected value of additional health. Because of the operational problems in implementing the above measures of need, many studies use self-reported indicators of morbidity as proxies for need for health care. In this paper, need for health care is proxied by a series of morbidity indicators. Self-reported health is captured by dummies: if the respondent reports her health to be fair or poor, good, or very good, respectively; the reference category is an excellent state of health. The NPHS data also has a generic health status index score, which is able to combine both quantitative and qualitative aspects of health. This index is based on the Comprehensive Health Status Measurement System developed at McMaster University's Centre for Health Economics and Policy Analysis. The index essentially provides a summary description of an individual's overall functional health, based on eight attributes: vision, hearing, speech, mobility, dexterity, cognition, emotion, and pain and discomfort. Since a higher scale indicates better health, we have used this index score as a measure of need as well. A derived number of chronic conditions based on more than 21 defined illness conditions[11] in the past 12 months is also used as an indicator of need. If there are no chronic conditions, the chronic variable is defined as 0. The number of disability days during the past two-weeks has also been used as an indicator of need.

Life-style Variables: The life style variables are represented by the extent of drinking, physical activity and smoking behaviour. Average daily alcohol consumption is a continuous variable. Frequency of all physical activity index is represented by two dummies: for moderate and inactive; the reference category for physical activity is active. Three dummy variables are used to examine the effect of smoking behaviour on utilization of health care: the respondent is a daily smoker, the respondent smokes occasionally, and a household member smoke inside the home (passive smoker). Table 3 explains

[11] The defined chronic illnesses are food allergies, allergies other than food allergies, asthma, arthritis or rheumatism, back problems excluding arthritis, high blood pressure, migraine headaches, chronic bronchitis or emphysema, sinusitis, diabetes, epilepsy, heart disease, cancer, stomach or intestinal ulcers, stroke, urinary incontinence, bowel disorder/Crohn's Disease or colitis, cataracts, Alzheimer's disease or other dementia, glaucoma, thyroid condition, and others.

the summary statistics of all explanatory variables used in this study.

6 Econometric Results and Analysis

In this section, we discuss the model specification results and provide a brief summary of our empirical findings. As indicated in the previous section, the natural starting point for count data models is to test for overdispersion. The likelihood ratio test statistics for overdispersion test results are presented in Table 4. From Table 4, it is clear that imposing the restriction that the conditional mean is equal to conditional variance (i.e., the Poisson regression model) is inappropriate. As stated earlier, the presence of overdispersion leads to inadequate model predictions; that is the predicted probabilities are incorrect, especially the number of zeros. Overdispersion in the data could arise either due to cross-sectional heterogeneity, a separate data generating process for 'zero-states', or specification errors.

Although the NB regression model performs much better than the Poisson model, we cannot rule out more than one stochastic process. In order to test the appropriateness of the assumption of two data generating processes and to test for the existence of cross-sectional heterogeneity, estimation of the

Poisson, ZIP, and ZINB models, and tests between them have been carried out. We present the test results in Table 5. It is quite evident that the test procedures favour ZINB over both ZIP and NB for all selected types of health care utilization.

As indicated earlier, the motivation for the hurdle or two-part model is the statistical representation of a principal-agent framework. One way to proceed with the hurdle model is to use either a zero-truncated Poisson or zero-truncated Negative Binomial model for positive outcomes. Since the zero truncated Poisson is nested in the zero-truncated Negative Binomial model, a likelihood ratio test for $H_0 : \alpha = 0$ is appropriate. It is evident that a zero-truncated negative binomial model is preferred to a zero-truncated Poisson model.

Now, we turn to discussion of the performance of three competing econometric specifications: the zero-inflated negative binomial model, the hurdle model, and a latent class model characterized by two latent classes. The specification of the zero-inflated negative binomial model consists of a probit model for the contact decision (zero outcome) and a negative binomial for the number of visits (non-zero outcomes). The specification of the hurdle model consists of a probit model for the contact decision and a truncated negative binomial for the intensity of utilization. The latent class model allows for a different constant term, different slope coefficients, and a different dispersion parameter where a NB model is identified by the data.

In order to compare the performance of these three models, we use the Log likelihood, Akaike Information Criterion (AIC = - ln L + 2K), and Bayesian Information Criterion (BIC = - 2ln L + K ln (N)) where ln L is the maximized log likelihood, K is the number of parameters, and N is the sample size. We prefer the model with bigger values of the log-likelihood and smaller values of AIC and BIC. We present these results in Table 6.

For the doctor and GP equations, all three criterias clearly favour latent class models. However, for the dentist and night equations, the hurdle model is preferred and for the specialist equation, the zero-inflated negative binomial model is preferred. It is to be noted that for the specialist, dentist and nights equations, we were not able to estimate NB specification for each class and the dispersion parameter across latent classes are not identifiable, perhaps due to over parametrization. The detailed results are presented in Table 7a through Table 9.

The effect of having prescription insurance on the utilization of doctor's services, GP's services and specialist's services are positive and statistically significant. Similarly, the effects of having dental insurance is positively significant for dentist's services. However, the effect of hospitalization insurance is either insignificant or negatively significant. This suggests that an increase in the probability of having supplemental insurance leads to an increase in *ex ante* demand for the services of GP, specialist and dentist. So, one might conclude that there is *ex ante* moral hazard in the utilization of health care through private supplemental health insurance for non-hospitalized services. However, in the context of latent class models, the insurance coefficients for class1 model are positive but insignificant for GP and doctor visits. This could mean that the effect of moral hazard is relatively lower for healthier groups in the utilization of GP's services.

There are indications of supplier-induced demand for physician visits, because the insurance coefficients for these services are positive and significant at the second stage of two stage models. However, the results are not significant for specialist visits. Another way to look into the supplier-induced demand for health care is to find out how the physician density affects health care utilization. In our model, this is captured through the geographical variables, since it is well known that physician density is highest in the metropolitan areas and relatively higher in the urban areas compared to that of the rural areas. We find that the geom variable is positively significant across all physician visits, which implies that physician density does have some impact on health care utilization.

Both the decision to contact and the decision to utilize GP services are responsive to need, proxied by morbidity indicators. The effects pertaining to the dummies for self reported health status and chronic illnesses are all positively significant at the 1% level. We can now test the hypothesis of horizontal and vertical equity in the utilization of GP services. According to Abasolo et al. [1], "horizontal equity requires that differential utilization of GP services between individuals should relate only to differences in their needs", while "vertical equity dictates that individuals with greater need make greater use of GP services". Thus, the null hypothesis of horizontal equity in utilization is: $H_0 : \frac{\partial GF}{\partial (Non{-}need)} = 0$, and vertical equity in utilization is: $H_0 : \frac{\partial GP}{\partial (Need)} > 0$, and $\frac{\partial GP}{\partial (Need)_2} > \frac{\partial GP}{\partial (Need)_1} > 0$, where $(Need)_2$ represents higher need. This is one interpretation of vertical equity adopted in empirical analysis [1]. Since our results suggest that greater utilization is associated with greater need for doctor visits, GP visits, specialist visits and nights spent in hospital, it can be interpreted as vertical equity in utilization of both hospitalized and non-hospitalized services in Canada. Further, this result is consistent with all econometric model specifications as well. For instance, the probability of being perceived fair/poor health, good health, and very good health leads to an increase in doctor visits by a factor of 1.32, .57, and .30, for class 2 and by a factor of .67, .43, and .23 for class 1 in relation to excellent health status, respectively.[12] The effect of chronic condition is positive and statistically significant for doctor visits as well as visits to GPs and specialists separately. The disability variable exhibits a similar trend. This result is again reinforced in the negatively significant hsiscore variable; the reporting of higher hsiscore (higher health status) leads to lower utilization of health services.

Family income appears to be an important determinant of non-hospitalized services including dentist visits, with higher income individuals tending to use more health care. However, the latent class model results suggest that income is a significant determinant of health care utilization for the relatively less healthy class.

Now, turning to the demographic variables, we find that gender affects utilization of all non-hospitalized services positively. This implies that women tend to seek more care than men, as evident in most empirical studies. Age

[12]The interpretation is based on exponential transformation of the coefficients.

dummies are generally insignificant or negatively significant for all physician visits and positive for nights spent in hospital for relatively higher age groups. The education variables are positively significant for specialist visits and dentist visits.

The life-style variables do not exhibit a clear sign and are generally insignificant. This seems rather strange. One possible explanation is that smoking, drinking and exercise habits do not immediately affect the realization of $s_{h_i} \in S_h$, but may affect the health status in the long-run leading to higher demand for health care in future. Since there is nothing in our dependent variable to capture this, these variables show irregular signs.

7 Conclusions

In this paper, we analyzed the underlying factors determining the utilization of different types of health care by using the most recent NPHS data. The number of visits to a health professional or the number of nights spent in hospital is used as a measure of utilization. We used a simple and intuitive microeconometric framework for analyzing the utilization of different types of health care. It is found that the decision to contact a health professional (i.e., *ex ante* utilization) and the decision about how much to utilize, proxied by the number of visits (i.e., *ex post* utilization), are essentially two distinct stochastic processes. The econometric results correspond to the underlying economic behaviours of two stochastic decision making processes relating to unobserved health improvements of the individuals. However, the latent class modeling framework suggests that it is a superior statistical technique if the data permits modeling unobserved heterogeneity and overdispersion.

Although many of our results are consistent with the literature, there are inherent limitations and potentials for improvements. As far as the limitations of this study are concerned, they may be at two levels, one at the data source and the other at the technical level. At the data source, the length of recalling period is subject to individual bias. Not every individual can remember and report every single incidence of illness episodes leading to the number of visits to different health professionals accurately. There may be time inconsistency in individual behaviours. This means that if the same individual is asked the same questions, he or she may not report the same answer. Since there is a great deal of subjectiveness in many questions, some people might be able to overstate the true state of affair while others may understate. However, in a large sample we might expect that there will be net out in the effects of those biases.

Important variables missing in the survey data are waiting time, travel time, and out of pocket spending for each visit to different types of health professionals. This constrains us to identify some of the interesting and crucial parameters relating to the demand for health care.

One technical improvement could be that we may have endogenous regressors and selection bias. One plausible endogeneity problem could be that the self reported health status indicators may itself be determined by the other regressors in the model. So a technical improvement is to account for the endogeneity of health status along the lines of Greene [17], Terza [45], and Windmeijer and Santos-Silva [50].

References

[1] Abasolo, I., Manning, R., and Jones, A. M. (2001), "Equity in utilization of and access to public sector GPs in Spain", *Applied Economics*, 33, 349-364.

[2] Arrow, K. J. (1963), "Uncertainty and the welfare economics of medical care", *American Economic Review*, **53**, 941-973.

[3] Cameron, A. C. Trivedi, P. K. (1986), "Econometric models based on count data: comparisons and implications of some estimators and tests", *Journal of Applied Econometrics*, **1,** 29-53.

[4] Cameron, A. C. Trivedi, P. K., Milne, F., and Piggot, J. (1988), "A microeconometric model of the demand for health care and health insurance in Australia", *Review of Economic Studies*, 55(1), 85-106.

[5] Cameron, A. C. Trivedi, P. K. (1990), "Regression based tests of overdispersion in the Poisson model", *Journal of Econometrics*, 46, 255-274.

[6] Cameron, A. C. Trivedi, P. K. (1998), "Regression Analysis of Count Data", Econometric Society Monograph 30, New York, Cambridge University Press.

[7] Cragg, J. G. (1971), "Some statistical models for limited dependent variables with applications to the demand for durable goods", *Econometrica* 39, 829-44.

[8] Culyer, A.J., Doorslaer, van E., and Wagstaff, A. (1992), "utilization as a measure of equity by Mooney, Hall, Donaldson and Gerard", *Journal of Health Economics*, 11, 93-98.

[9] Davidson, J., and Mackinon, J. (1993), "Estimation and Inference in Econometrics", New York, *Oxford University Press.*

[10] Dev and Trivedi (1997), "Demand for medical care by the elderly: a finite mixture approach", *Journal of Applied Econometrics* 12, 313-336.

[11] Dev, 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.

[12] Doorslaer, van E., and Wagstaff, A. (1992), "Equity in the delivery of health care: some international comparisons", *Journal of Health Economics*, 11, 389-411.

[13] Erbsalnd, M. Reid, W. Ulrich, V. Health, health care, and the environment. Econometric evidence from German micro data. *Health Economics* 1995; 4: 169-182.

[14] Gerdtham, U.-G. (1997), "Equity in health care utilization: further tests based on hurdle models and Swedish micro data", *Health Economics*, 6, 303-319.

[15] Gerdtham U.-G. Johanesson, M. Lundberg, L. Isacson, D. The demand for health: results from new measures of health capital. *European Journal of Political Economy* 1999; 15(3): 501-521.

[16] Gerdtham, U.-G. Johanesson, M. New estimates of the demand for health: results based on a categorical health measure and Swedish micro data. *Social Science & Medicine* 1999; 49: *1325-1332.*

[17] Greene, W. H. (1994), "Accounting for excess zeros and sample selection in Poisson and Negative Binomial models", *Stern School of Business Working Papers*, Ec-94-10, New York University.

[18] Grogger, J. T. and Carson, R. T. (1991), "Models for truncated counts", *Journal of Applied Econometrics*, 6, 225-238.

[19] Grossman, M. (1972a), "On the concept of health capital and the demand for health", *Journal of Political Economy*, 80(2), 223-235.

[20] Grossman, M. (1972b), "The demand for health - a theoretical and empirical investigation", (NBER) New York, Columbia University Press.

[21] Grossman, M. (2000), "The human capital model", in J. P. Newhouse and A. J. Culyer, eds. Handbook of Health Economics, volume 1a, Elsvier, chapter 7.

[22] Gurmu, S. and Trivedi, P. K. (1996), ""Excess zeros in count models for recreational trips, *Journal of Business and Economic Statistics*, 14, 469-477.

[23] Gurmu, S. (1997), "Semi-parametric estimation of hurdle regression models with an application to Medicaid utilization", *Journal of Applied Econometrics*, 12, 225-242.

[24] Gurmu, S. (1998), "Generalized hurdle count data regression models", *Economic Letters*, 58, 263 268.

[25] Hausman, J. A.. Hall, B. H., and Griliches , Z. (1984), "Econometric models for count data with an application to the Patents-R and D relationship", *Econometrica*, 52, 909-938.

[26] Heckman, J. J. and Singer, B. A. (1984), "A method of minimizing the distributional impact in econometric models for duration data", *Econometrica*, 52, 271-320.

[27] Heckman, J. J. (2000), "Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture", *Journal of Political Economy*, 109(4), 673-748.

[28] Johnson, N. L., and Kotz, S. (1969), "Discrete Distributions", Boston, Houghton Mifflin.

[29] Jones, A. M. (2000), "Health econometrics", in J. P. Newhouse and A. J. Culyer, eds., Handbook of Health Economics, volume 1a, Elsevier, 265-364.

[30] Lambart, D. (1992), "Zero-inflated Poisson regression with an application to defects in manufacturing", *Technometrics*, 34, 1-14.

[31] MaCullagh, P., and Nelder, J. A. (1983), "Generalised Linear Models", Chapman and Hall, London.

[32] Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., Leibowitz, A., Marquis, M. S., and Phelps, C. E. (1981), "A two part model of the demand for medical care: preliminary results from the health insurance study." in J. van der Gaag and M. Perlman, eds., *Health Economics, and Health Economics*, Amsterdam, North Holland, 101-123.

[33] Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., Leibowitz, A., and Marquis, M. S. (1987), "Health insurance and the demand for medical care: evidence from a randomized experiment", *American Economic Review*, **77**, 251-277.

[34] Manning, W. G., Duan, N. and Rogers, W. H. (1987), "Monte carlo evidence on the choice between sample selection and two-part models", *Journal of Econometrics*, 35, 59-82.

[35] Manning, W. G. and Mullahy, J. (2001), "Estimating log models: to transform or not to transform?", *Journal of Health Economics*, 20, 461-494.

[36] Mullahy, J. (1986) "Specification and testing of some modified count data models", *Journal of Econometrics*, **33**, 341-65.

[37] Mullahy, J. (1997), "Heterogeneity, excess zeros, and the structure of count data models", *Journal of Econometrics*, 12, 337-350.

[38] Mullahy, J. (1998), "Much ado about two: reconsidering retransformation and the two-part model in health econometrics", *Journal of Health Economics*, 17, 247-281.

[39] Muurinen, J. M. (1982), "Demand for health: a generalised Grossman model", *Journal of Health Economics*, 1(1), 3-28.

[40] Newbold, K.B., Eyles, J., and Birch, S. (1995), "Equity in health care: methodological contributions to the analysis of hospital utilization within Canada", "*Social Science and Medicine*", **40(9)**, 1181-1192.

[41] Pohlmeier, W. and Ulrich, V. (1995), "An econometric model of the two part decision making process in the demand for health care", *The Journal of Human Resources*, XXX (2), 339-361.

[42] Santos-Silva, J. M. C., and Windmeijer, F. (2001), "Two-part multiple spell models for health care demand", *Journal of Econometrics*, 104, 67-89.

[43] Stabile, M. (2001), "Private insurance subsidies and public health care markets: evidence from Canada", *Canadian Journal of Economics*, 34(4), 921-942.

[44] Stoddart, G. L. and Barer, M. L. (1981), "Analysis of the demand and utilization through episodes of medical care", in J. van der Gaag and M. Perlman, eds., *Health Economics, and Health Economics*, Amsterdam, North Holland.

[45] Terza, J. V. (1997), "Estimating count data models with endogenous switching: sample selection and endogenous treatment effects", *Journal of Econometrics*, 84, 129-154.

[46] Vuong , Q. (1989), "Likelihood ratio tests for model selection and non-nested hypotheses", *Econometrica*, 57, 307-334.

[47] Wagstaff, A. (1986), "The demand for health: some new empirical evidence", *Journal of Health Economics*, 7(2), 151-163.

[48] Wagstaff, A., Doorslaer, van E., and Paci, P. (1991), "On the measurement of horizontal equity in the delivery of health care", *Journal of Health Economics*, 10, 169-205.

[49] Wagstaff, A. (1993), "The demand for health: an empirical reformulation of the Grossman model", *Health Economics*, **2**, 189-198.

[50] Windmeijer, F. A. G. and Santos-Silva, J. M. C. (1997), "Endogeneity in count data models: an application to demand for health care", *Journal of Applied Econometrics*, 12(3), 281-294.

[51] Winkelmann, R. and K. F. Zimmermann (1995), 'Recent developments in count data modelling: theory and application', *Journal of Economic Surveys*, 9, 1-24.

[52] Zweifel, P. (1981), "Supplier induced demand in a model of physician behaviour", in J. van der Gaag and M. Perlman, eds., *Health Economics, and Health Economics*, Amsterdam, North Holland, 245-267.

[53] Zweifel, P. Bryer, F. Health Economics. Oxford University Press, 1997.

[54] Zweifel, P. and Manning, W. G. (2000), "Moral hazard and consumer incentives in health care", in J. P. Newhouse and A. J. Culyer, eds. Handbook of Health Economics, volume 1a, Elsvier, chapter 8.

Table1
Variable Definitions

Doctor	Number of consultations in 12 months to Doctors (GP + Specialists).
GP	Number of consultations in 12 months to family doctor/general practitioner.
Specialist	Number of consultations in 12 months to Other specialized medical doctor (such as allergist, gynecologist, or psychiatrist).
Dentist	Number of consultations in 12 months to Dentist or orthodontist.
Nights	Number of nights spent as patient during past 12 months.
female	male = 0, female = 1.
single	single = 1, otherwise = 0.
married	married = 1, otherwise = 0.
wsd	widow, separated, and divorced = 1, otherwise = 0.
age18	Age group: 15 - 34 years.
age28	Age group: 35 - 59 years.
age38	Age group: 60 - 74 years.
age48	Age group: 75 or older.
yimmig0	Immigrant: no = 1, yes = 0.
yimmig1	Years of immigration: 0 - 4 years.
yimmig2	Years of immigration: 5 - 9 years.
yimmig3	Years of immigration: 10 or more years.
geor	rural area = 1, otherwise = 0.
geou	urban area = 1, otherwise = 0.
geom	metropolitan area = 1, otherwise = 0.
edu_lsc	Education: less than secondary = 1, otherwise = 0.
edu_sec	Education: completed secondary = 1, otherwise = 0.
edu_grd	Education: completed post-secondary = 1, otherwise = 0.
inc_liq	Income: lower income quintile = 1, otherwise = 0.
inc_miq	Income: middle income quintile = 1, otherwise = 0.
inc_umiq	Income: upper middle income quintile = 1, otherwise = 0.
inc_hiq	Income: high income quintile = 1, otherwise = 0.
ins_prs	Insurance: Prescription medication = 1, otherwise = 0.
ins_den	Insurance: Dental = 1, otherwise = 0.
ins_hosp	Insurance: Hospital charges=1, otherwise = 0.
fair/poorh	Health Status: Fair/Poor = 1, otherwise = 0.
goodh	Health Status: Good = 1, otherwise = 0.
vgoodh	Health Status: Very Good = 1, otherwise = 0.
excellh	Health Status: Excellent = 1, otherwise = 0.
pai_act	Physical activity index: active = 1, otherwise = 0.
pai_mod	Physical activity index: moderate = 1, otherwise = 0.
pai_inc	Physical activity index: inactive = 1, otherwise = 0.
hsiscore	Health Utility Index.
chronic	Number of chronic conditions.
disability	Total number of disability days during the past two weeks.
alcddly	Average daily alcohol consumption.
smoker	Smoke: yes = 1, no = 0.
smk_dly	Smoke daily = 1, otherwise = 0.
smk_occ	Smoke occassionally = 1, otherwise = 0.
hhsmoke	Family member(s) smoke inside house: yes = 1, no = 0.
pr_nfld	Province: Newfoundland = 1, otherwise = 0.
pr_pei	Province: Prince Edward Island = 1, otherwise = 0.
pr_ns	Province: Nova Scotia = 1, otherwise = 0.
pr_nb	Province: New Brunswick = 1, otherwise = 0.
pr_que	Province: Qu´ebec = 1, otherwise = 0.
pr_ont	Province: Ontario = 1, otherwise = 0.
pr_mb	Province: Manitoba = 1, otherwise = 0.
pr_sask	Province: Newfoundland = 1, otherwise = 0.
pr_ab	Province: Saskatchewan = 1, otherwise = 0.
pr_bc	Province: British Columbia = 1, otherwise = 0.

Table 2a
Frequency Distribution of Consultations

Number of Consultations	Doctor	GP	Specialist	Dentist	Nights
0	1820	2030	7300	4090	9059
1	1905	2169	1113	2666	191
2	1630	1775	562	1987	105
3	1040	1034	272	478	100
4	770	809	178	286	58
5	508	379	82	107	48
6	452	491	78	75	23
7	255	101	33	19	42
8	215	146	32	25	17
9	122	36	12	5	11
10	174	188	43	19	28
11	83	11	4	-	5
12	244	332	84	36	12
13	85	7	-	-	6
14	66	20	-	-	26
15	67	62	-	-	9
16	54	7	-	-	4
17	23	8	-	-	4
18	30	9	-	-	3
19	9	2	-	-	-
20	36	59	-	-	3
21	10	1	-	-	4
22	16	3	-	-	2
23	7	2	-	-	-
24	18	35	-	-	-
25	15	14	-	-	-
26	21	9	-	-	-
27	9	-	-	-	1
28	15	1	-	-	2
29	2	-	-	-	-
30	11	18	-	-	10
31+	81	35	-	-	20
Total	9793	9793	9793	9793	9793

Table 2b
Characteristics of the Dependent Variables

Variables	Mean	Variance	Skewness	Kurtosis	Min	Max
Doctor	3.896559	26.23164	2.723406	12.19107	0	31
GP	3.188094	18.71339	3.096233	15.67967	0	31
Specialist	.6862044	3.10669	4.020823	21.90118	0	12
Dentist	1.14398	2.253982	2.810081	16.317	0	12
Nights	.4363321	5.873014	8.675457	92.8232	0	31

Table 3
Summary Statistics for Explanatory Variables

Variable	Obs	Mean	Std. Dev.	Min	Max
female	9793	.5192484	.4996549	0	1
married	9793	.5787808	.4937798	0	1
single	9793	.2509956	.4336081	0	1
wsd	9793	.1702236	.3758483	0	1
age1	9793	.331257	.4706893	0	1
age2	9793	.4693148	.499083	0	1
age3	9793	.1368324	.3436879	0	1
age4	9793	.0535076	.2250549	0	1
yimmig0	9793	.8693965	.3369834	0	1
yimmig1	9793	.0122537	.1100215	0	1
yimmig2	9793	.0130706	.1135827	0	1
yimmig3	9793	.1052793	.3069286	0	1
geor	9793	.2198509	.4141666	0	1
geou	9793	.5773512	.4940058	0	1
geom	9793	.2027979	.4021038	0	1
edu_lsc	9793	.226182	.4183797	0	1
edu_sec	9793	.4143776	.4926394	0	1
edu_grd	9793	.3594404	.4798609	0	1
inc_liq	9793	.1375472	.3444417	0	1
inc_miq	9793	.261207	.4393149	0	1
inc_umq	9793	.3868069	.4870437	0	1
inc_hiq	9793	.2144389	.4104535	0	1
ins_prs	9793	.723476	.4473018	0	1
ins_den	9793	.5611151	.4962762	0	1
ins_hsp	9793	.6089043	.4880207	0	1
fpoorh	9793	.0802614	.2717113	0	1
goodh	9793	.2535485	.4350643	0	1
vgoodh	9793	.4243848	.4942745	0	1
excellh	9793	.2418054	.4281989	0	1
pai_act	9793	.2195446	.4139592	0	1
pai_mod	9793	.2564076	.4366718	0	1
pai_inc	9793	.5240478	.4994469	0	1
hsiscore	9793	.9019413	.1660775	-.174	1
chronic	9793	1.296743	1.520411	0	10
disability	9793	.7704483	2.646561	0	14
alcddly	9793	.5105688	.9691502	0	14
hhdsmok	9793	.3435107	.4749043	0	1
smoker	9793	.3199224	.46647	0	1
smk_dly	9793	.2768304	.4474548	0	1
smk_occ	9793	.043092	.2030746	0	1
pr_nfld	9793	.0575922	.2329825	0	1
pr_pei	9793	.0510569	.220125	0	1
pr_ns	9793	.0614725	.2402072	0	1
pr_nb	9793	.0576943	.2331763	0	1
pr_que	9793	.1782906	.3827767	0	1
pr_ont	9793	.2717247	.4448714	0	1
pr_mb	9793	.0698458	.2549	0	1
pr_sask	9793	.0625957	.2422468	0	1
pr_ab	9793	.0937404	.2914821	0	1
pr_bc	9793	.0959869	.2945884	0	1

Table 4
Overdispersion Test Results

Model	NB1: $H_0 = \alpha_1 = 0$		NB2: $H_0 = \alpha_2 = 0$	
	Estimate of α_1	LR Statistics	Estimate of α_2	LR Statistics
Doctor	2.902147	15000	.7396861	15000
GP	2.244423	11000	.68142	11000
Specialist	3.022224	7612.07	3.895602	7312.29
Dentist	.6298126	1540.68	.4770391	1238.29
Nights	14.40082	15000	24.99612	11000

Table 5
Test Between Zero and Non-Zero Inflated Models

Model	ZIP vs. Poisson	ZINB vs. NB	ZINB vs. ZIP
	Vuong Test	Vuong Test	LR Statistics
Doctor	18.89	6.84	11000
GP	16.37	6.35	8390.68
Specialist	20.80	11.51	1671.24
Dentist	12.76	11.33	568.67
Nights	17.48	7.93	1695.23

Table 6
Model Specification Testing

Model	NB1	NB2	ZINB (Probit)	Hurdle (Probit)	Latent Class
Doctor					
Log Likelihood	-22938.776	-22903.565	-22795.64	-22664.45	-22586.39+[a]
AIC	45957.55	45887.13	45749.28	45486.9	45336.78[b]
BIC	46245.13	46174.71	46317.24	46054.86	45926.31[c]
GP					
Log Likelihood	-21300.776	-21187.522	-21095.44	-20933.79	-20845.45+[a]
AIC	42681.55	42455.04	42348.88	42025.58	41854.9[b]
BIC	42969.13	42742.62	42916.84	42593.54	42444.43[c]
Specialist					
Log Likelihood	-9375.1703	-9525.0616	-9298.992[a]	-9310.809	-9620.985++
AIC	18830.34	19130.12	18755.98[b]	18779.62	19401.97
BIC	19117.92[c]	19417.7	19323.95[d]	19347.58	19977.12
Dentist					
Log Likelihood	-13757.248	-13908.442	-13649.93	-13502.80[a]	-13564.92++
AIC	27594.5	27896.88	27457.86	27163.6[b]	27289.94
BIC	27882.07	28184.46	28025.82	27731.56[c]	27864.99
Nights					
Log Likelihood	-4310.0066	-4352.128	-4214.378	-4213.389[a]	-4778.520++
AIC	8700.013	8784.256	8586.756	8584.778[b]	9717.04
BIC	8987.59[c]	9071.833	9154.72	9152.742[d]	10292.19

[a] Model with the bigger log likelihood value; [b] Model preferred by the AIC; and [c] Model preferred by the BIC; [d] Should be preferred by the BIC because NB model was rejected previously.
+ Latent class probabilities are modeled as Negative Binomial model
++ Latent class probabilities are modeled as Poisson model. Latent class Negative Binomial model seems to be overparameterized for these data sets.

Table 7a
Utilization of Doctor's Services

Variables	NB2	Hurdle Model		ZINB-Probit		Latent Class	
		Stage 1 Probit	Stage 2 ZTNB	Zero Outcome	Non Zero Outcome	Class 1	Class 2
constant	1.0458* (11.67)	.2301 (1.441)	1.1406* (10.095)	-.8151 (-1.21)	1.1645* (12.95)	2.2259* (8.582)	.1732 (1.439)
female	.345* (15.5)	.4738* (13.932)	.2641* (9.980)	-4.423 (-0.30)	.2672* (11.52)	.2436* (4.687)	.3741* (13.056)
married	.1042* (3.67)	.104** (2.444)	.0875* (2.754)	-.1371 (-.97)	.0849* (2.93)	.1274* (2.067)	.0265 (.728)
wsd	.042 (1.11)	-.014 (-.240)	.0524 (1.248)	.3716 (-1.39)	.029 (.77)	.0986 (1.239)	-.041 (-.870)
age28	-.1562* (-5.84)	-.0656*** (-1.632)	-.1988* (-6.850)	-.263*** (-1.77)	-.1798* (-6.58)	-.2284* (-3.908)	-.045 (-1.287)
age38	-.1752* (-4.57)	.0304 (.487)	-.2501* (-5.194)	-.5913*** (-1.79)	-.209* (-5.45)	-.4272* (-4.381)	.0745 (1.416)
age48	-.0829 (-1.56)	.0865 (.918)	-.1437** (-2.332)	-.6087 (-1.05)	-.1169** (-2.22)	-.2082*** (-1.704)	.0538 (.826)
yimmig1	-.0839 (-.85)	-.0014 (-.011)	-.101 (-.759)	-4.5209 (-.06)	-.1357 (-1.42)	-.3904 (-1.445)	.144 (1.121)
yimmig2	-.0733 (-.77)	.0277 (.207)	-.0856 (-.725)	-.9184 (-.64)	-.1077 (-1.05)	-.4421** (-2.33)	.2631** (2.058)
yimmig3	.0455 (1.28)	.1589* (2.730)	.0077 (.185)	-.8605 (-1.40)	.0172 (.48)	-.0512 (-.618)	.1315* (2.928)
geou	.0417 (1.53)	.0886** (2.193)	.0216 (.661)	-.0722 (-.52)	.0358 (1.30)	.0404 (.646)	.037 (1.076)
geom	.162* (4.38)	.1244* (2.256)	.1563* (3.695)	-.1868 (-.87)	.1416* (3.79)	.2524* (3.126)	.047 (.983)
edu_sec	-.0412 (-1.44)	.0185 (.425)	-.0597*** (-1.734)	-.1651 (-1.10)	-.048*** (-1.66)	-.094 (-1.375)	.0226 (.636)
edu_grd	.0196 (.64)	.1174** (2.490)	-.0157 (-.434)	-.1648 (-.96)	.0088 (.28)	-.0009 (.012)	.0609 (1.607)
inc_miq	.003 (.09)	.0876*** (1.631)	-.0272 (-.652)	-.3135*** (-1.64)	-.0072 (-.21)	-.0374 (-.455)	.0813*** (1.883)
inc_umiq	-.0091 (-.26)	.1485* (2.766)	-.0641 (-1.550)	-.2607 (-1.42)	-.0142 (-.40)	-.1306 (-1.609)	.1456* (3.215)
inc_hiq	-.0169 (-.26)	.1695* (2.749)	-.084*** (-1.737)	-.4542** (-2.05)	-.0328 (-.80)	-.172*** (-1.748)	.1493* (2.805)
ins_prs	.0927* (3.76)	.1374* (3.682)	.0626** (2.242)	-.2274*** (-1.70)	.0766* (3.06)	.0752 (1.410)	.1113* (3.575)
fair/poorh	.6957* (14.48)	.3542* (3.976)	.7735* (12.544)	.0757 (.19)	.7119* (14.92)	.5071* (4.2)	.842* (13.999)
goodh	.4083* (12.75)	.2004* (4.279)	.4703* (12.680)	.3933** (2.04)	.427* (13.06)	.363* (4.978)	.4546* (10.753)
vgoodh	.241* (8.58)	.1436* (3.752)	.2679* (8.763)	.285*** (1.73)	.2545* (8.82)	.213* (3.527)	.265* (6.910)
hsiscore	-.5995* (-8.43)	-.2602*** (-1.808)	-.6557* (-6.979)	.3689 (.60)	-.5723* (-8.20)	-.6477* (-3.347)	-.4178* (-5.145)
chronic	.1828* (22.61)	.273* (18.544)	.1669* (17.070)	-1.1351* (-6.03)	.1648* (20.7)	.1357* (6.834)	.1998* (20.896)
disability	.0399* (10.33)	.0642* (6.262)	.0382* (7.094)	-.2586 (-1.24)	.0381* (10.2)	.0151 (1.317)	.055* (11.428)
alcddly	-.0015	-.0283***	.0171	.102**	.0149	.016	-.1715

	(-.13)	(-1.798)	(1.373)	(2.37)	(1.19)	(.713)	(-1.174)
pai_mod	-.0388 (-1.26)	-.0165 (-.361)	-.034 (-.968)	.4052** (2.21)	-.0128 (-.41)	-.4565 (-.653)	-.0411 (-1.052)
pai_inc	-.0002 (-.01)	-.007 (-.163)	.0119 (.378)	.3047*** (1.73)	.0172 (.62)	-.0191 (-.312)	-.0174 (-.509)
smk_dly	-.0834* (-2.52)	-.133* (-2.735)	-.0545 (-1.431)	.2075 (1.26)	-.0656** (-1.95)	-.0058 (-.078)	-.1573* (-3.66)
smk_occ	.0827 (1.58)	.1008 (1.233)	.066 (1.154)	-.2299 (-.77)	.0687 (1.30)	.104 (.946)	.049 (.727)
hhsmoke	-.035 (-1.15)	-.0334 (-.727)	-.034 (-.958)	-.0747 (-.47)	-.041 (-1.33)	-.074 (-1.067)	.016 (.404)
pr_nfld	.0877*** (1.77)	.136*** (1.762)	.0585 (.964)	-.4442 (-1.54)	.0638 (1.27)	-.1387 (-1.204)	.2718* (4.084)
pr_pei	.0797 (1.55)	.2615* (3.017)	.0002 (.004)	-.8652** (-1.95)	.0388 (.76)	-.057 (-.449)	.2072* (3.138)
pr_ns	.076 (1.60)	-.0513 (-.691)	.1056*** (1.896)	.2296 (.97)	.0838*** (1.76)	.015 (.132)	.0999*** (1.736)
pr_nb	-.0926*** (-1.86)	-.042 (-.544)	-.1047*** (-1.816)	.1035 (.43)	-.0811*** (-1.61)	-.1233 (-1.045)	-.068 (-1.082)
pr_que	-.2183* (-6.51)	-.1799* (-3.639)	-.2213* (-5.942)	.0923 (.51)	-.2107* (-6.17)	-.3303* (-4.603)	-.1007** (-2.274)
pr_mb	-.0783*** (-1.71)	-.0914 (-1.333)	-.0825 (-1.528)	.1013 (-.42)	-.0854*** (-1.85)	-.1211 (-1.160)	-.041 (-.690)
pr_sask	.0284 (.60)	.032 (.437)	.0181 (.342)	-.269 (-.95)	.0106 (.22)	-.0312 (-.312)	-.0905 (1.551)
pr_ab	-.088 (-2.15)	-.161** (-2.621)	-.0663 (-1.344)	.2777 (1.26)	-.0766*** (-1.84)	-.1635*** (-1.742)	-.0447 (-.861)
pr_bc	-.0492 (-1.25)	-.115*** (-1.890)	-.031 (-.644)	-.321 (-1.01)	-.0615 (-1.57)	-.1974** (-2.1)	-.0574 (1.161)
-Log L	22903.565	22664.45		22795.64		22586	
Alpha (α)	.739686	1.0533 (25.465)		.6778		1.6885* (4.869)	2.9558* (12.391)
Class Prob						.2318* (5.289)	.7682* (17.529)

Figures in parenthesis are t-statistics.
The symbols *, **, and *** denote 1%, 5%, and 10% significance, respectively.

Table 7b
Utilization of GP's Services

Variables	NB2	Hurdle Model		ZINB-Probit		Latent Class	
		Stage 1 Probit	Stage 2 ZTNB	Zero Outcome	Non Zero Outcome	Class 1	Class 2
constant	.8002* (8.98)	.2302 (1.556)	.8003* (7.054)	-.6419 (-0.83)	.9116* (10.07)	1.718* (6.778)	.1738 (1.608)
female	.2982* (13.35)	.40003* (12.253)	.2312* (8.348)	-4.6215 (-.23)	.2336* (10.07)	.2701* (4.677)	.3034* (10.781)
married	.0956* (3.34)	.1345* (3.299)	.0619*** (1.79)	-.2802*** (-1.70)	.0706 (2.40)	.1315*** (1.806)	.024 (.660)
wsd	.0435 (1.16)	.0389 (.696)	.0347 (.798)	-.5157*** (-1.71)	.0241 (.63)	.0859 (.983)	-.024 (-.523)
age28	-.1653* (-6.12)	-.0784** (-2.031)	-.2177* (-7.037)	-.2159 (-1.25)	-.1833* (-6.61)	-.2333* (3.525)	-.074** (-2.142)
age38	-.1563* (-4.08)	-.0099 (-.169)	-.2234* (-4.567)	-.688 (-1.56)	-.1858* (-4.81)	-.3471* (-3.311)	.155 (.322)
age48	-.0207 (-.39)	.0648 (.717)	-.6342 (-1.008)	-.4788 (-.80)	-.0483 (-.92)	-.0701 (-.520)	.0411 (.673)
yimmig1	-.033 (-.33)	.0189 (.148)	-.0396 (-.289)	-4.735 (-.04)	-.0785 (-.81)	-.2375 (-.843)	.085 (.639)
yimmig2	.0437 (.46)	.0117 (.148)	-.0967 (.819)	-.9808 (-.51)	.013 (.12)	-.1778 (-.829)	.2566* (2.197)
yimmig3	.0677*** (1.89)	.1497* (2.693)	.0411 (.926)	-1.6377 (-1.50)	.0396 (1.12)	-.036 (-.377)	.142* (3.251)
geou	.0322 (1.18)	.0649*** (1.653)	.0158 (.470)	-.0789 (-.50)	.0254 (.91)	.0395 (.586)	.025 (.733)
geom	.149* (3.99)	.0725 (1.366)	.171* (3.882)	-.1189 (-.43)	.1324* (3.49)	.2651* (2.944)	.0482 (1.018)
edu_sec	-.07** (-2.44)	-.0324 (-.771)	-.084** (-2.347)	-.1736 (-1.03)	-.0759* (-2.60)	-.1254*** (-1.686)	-.027 (-.780)
edu_grd	-.0477 (-1.56)	.0431 (.946)	-.0855** (-2.333)	-.2232 (-1.15)	-.0603** (-1.94)	-.0688 (-.905)	-.0085 (-.229)
inc_miq	.0038 (.11)	.0838 (1.608)	-.034 (-.797)	-.4374** (-2.08)	-.01 (-.28)	-.0585 (-.650)	.0956** (2.287)
inc_umiq	-.019 (-.54)	.1508* (2.901)	-.0937** (-2.212)	-.3433*** (-1.76)	-.0282 (-.79)	-.1584*** (-1.792)	.1372* (3.150)
inc_hiq	-.104* (-2.53)	.1494** (2.509)	-.2223* (-4.435)	-.6847* (-2.65)	-.1266* (-3.05)	-.348* (-3.291)	.1231** (2.318)
ins_prs	.0844* (3.41)	.1285* (3.564)	.0556*** (1.872)	-.256*** (-1.73)	.0692* (2.75)	.0708 (1.177)	.099* (3.247)
fair/poorh	.6806* (14.24)	.3261* (3.890)	.7946* (12.725)	.0481 (.10)	.6985* (14.64)	.5114* (4.055)	.8201* (14.148)
goodh	.399* (12.4)	.1941* (4.283)	.4964* (12.848)	.6099* (2.61)	.4247* (12.90)	.411* (5.075)	.4040* (9.510)
vgoodh	.2471* (8.7)	.134* (3.602)	.3055* (9.371)	.4266** (2.03)	.2658* (9.08)	.2801* (4.114)	.2309* (5.917)
hsiscore	-.4857* (-6.87)	-.1773 (-1.384)	-.5567* (-5.976)	.0222 (.03)	-.4703* (-6.71)	-.5176* (-2.658)	-.4101* (-5.3)
chronic	.1816* (22.73)	.2393* (17.682)	.1749* (17.728)	-1.0916* (-6.03)	.1663* (21.04)	.1786* (8.229)	.1753* (20.122)
disability	.0397* (10.42)	.0611* (6.650)	.0383 (7.381)	-.3058 (-1.28)	.038* (10.24)	.0259** (2.281)	.0466* (11.297)
alcddly	-.0014	-.03**	.0203	.1372*	.0173	.022	-.0213

	(-.12)	(-1.960)	(1.578)	(2.98)	(1.36)	(.877)	(-1.452)
pai_mod	-.0412 (-1.33)	-.0299 (-.679)	-.0342 (-.918)	.4169*** (1.91)	-.02 (-.64)	-.0693 (-.870)	-.0122 (-.311)
pai_inc	.003 (.12)	-.0024 (-.062)	.0163 (.487)	.348*** (1.64)	.0189 (.68)	-.0418 (-.590)	.0235 (.681)
smk_dly	-.0782** (-2.36)	-.146* (-3.078)	-.0382 (-.968)	.2346 (1.26)	-.0607*** (-1.80)	-.0264 (-.317)	-.1804* (-4.175)
smk_occ	.091*** (1.73)	.1157 (1.458)	.0821 (1.425)	-.367 (-1.00)	.0742 (1.41)	.1445 (1.228)	.0518 (.785)
hhsmoke	-.0253 (-.83)	-.0366 (-.818)	-.0226 (-.609)	-.1936 (-1.10)	-.0371 (-1.20)	-.1053 (-1.359)	.0537 (1.339)
pr_nfld	.1908* (3.87)	.093 (1.256)	.2195* (3.495)	-.3281 (-.96)	.1737* (3.47)	-.0047 (-.038)	.3356* (5.347)
pr_pei	.0941*** (1.82)	.2504* (3.007)	.1827 (.279)	-.6649 (-1.45)	.0645 (1.24)	-.4115 (-.287)	.2168* (3.306)
pr_ns	.1688* (3.59)	-.0353 (-.489)	.2415* (4.304)	.4619*** (1.68)	.1851* (3.88)	.1488 (1.244)	.153* (2.722)
pr_nb	-.0512 (-1.03)	-.0482 (.651)	-.0443 (-.770)	.2153 (.74)	-.0352 (-.69)	-.0435 (-.355)	-.0572 (-.915)
pr_que	-.2649* (-7.76)	-.2955* (-6.296)	-.231* (-5.916)	.2227 (.92)	-.255* (-7.21)	-.276* (-3.412)	-.2376 (-5.403)
pr_mb	-.0096 (-.21)	-.1187*** (-1.787)	.0324 (.582)	.0972 (.34)	-.0084 (-.18)	-.0095 (-.081)	-.017 (-.299)
pr_sask	.1313* (2.78)	.0455 (.635)	.1641* (2.901)	-.1437 (-.43)	.1199* (2.51)	.1127 (.969)	.152* (2.652)
pr_ab	-.0116 (-.28)	-.195* (-3.297)	.0682 (1.317)	.6155** (2.45)	.014 (.33)	-.0397 (-.377)	-.0113 (-.221)
pr_bc	.0545 (1.4)	-.0641 (-1.072)	.1005** (2.044)	-.3478 (-.82)	.0452 (1.15)	-.0492 (-.489)	.1144** (2.396)
-Log L	21187.522	20933.79		21095.44		20845.45	
Alpha (α)	.68142	1.0775 (22.971)		.6323		1.323* (6.535)	4.0482* (8.458)
Class Prob						.243* (6.335)	.757* (19.769)

Figures in parenthesis are t-statistics.
The symbols *, **, and *** denote 1%, 5%, and 10% significance, respectively.

Table 7c
Utilization of Specialist's Services

Variables	NB1	Hurdle Model		ZINB-Probit		Latent Class	
		Stage 1 Probit	Stage 2 ZTNB	Zero Outcome	Non Zero Outcome	Class 1	Class 2
constant	-1.33*	-1.3051*	.4234	.2904	.1747	1.2969*	-3.3309*
	(-8.19)	(-10.380)	(1.264)	(.68)	(.73)	(11.211)	(-17.353)
female	.4828	.3631*	.0566	-.8063*	.0099	.2488*	.758*
	(11.05)	(11.572)	(.711)	(-8.08)	(.15)	(7.781)	(13.574)
married	.001	.0083	.105	-.0467	.0186	.035	-.0892
	(.02)	(.202)	(1.134)	(-.52)	(.24)	(.995)	(-1.339)
wsd	-.067	-.056	.061	-.013	-.0402	-.0605	-.198**
	(-.93)	(-1.060)	(.516)	(-.09)	(-.42)	(-1.310)	(-2.394)
age28	.065	.05243	-.2217**	-.276*	-.2444*	-.1245*	.1382*
	(1.24)	(1.371)	(-2.513)	(-3.07)	(-3.26)	(-3.679)	(2.165)
age38	.075	.098***	-.5286*	-.5312*	-.4688*	-.3544*	.1732**
	(1.06)	(1.834)	(-4.237)	(-3.36)	(-4.60)	(-6.614)	(2.112)
age48	-.23**	-.102	-.5441*	-.717*	-.7185*	-.538*	-.5207*
	(-2.22)	(-1.404)	(-2.955)	(-2.05)	(-5.13)	(-7.243)	(-3.910)
yimmig1	-.2445	-.186	.203	.477*	.3252	-.0969	-.476
	(-1.1)	(-1.20)	(.538)	(1.84)	(1.00)	(-.861)	(-1.316)
yimmig2	-.1455	-.0896	-.8565**	-.114	-.5698***	-.7933*	-.204
	(-.71)	(-.622)	(-1.981)	(-.37)	(-1.89)	(-3.3)	(-.666)
yimmig3	.0207	.005	.0207	-.0343	.0022	.01	.007
	(.32)	(.104)	(.180)	(-.29)	(.02)	(.215)	(.102)
geou	.0631	.041	.0866	-.0107	.104	.019	.0826
	(1.2)	(1.074)	(.881)	(-.11)	(1.39)	(.476)	(1.268)
geom	.038	.019	.1856	.1314	.2045**	.1717*	.072
	(.55)	(.374)	(1.567)	(1.05)	(2.14)	(3.420)	(.923)
edu_sec	.2187*	.1818*	-.125	-.3181*	.0094	.0185	.3819*
	(3.83)	(4.521)	(-1.248)	(-3.33)	(.12)	(.491)	(5.549)
edu_grd	.358*	.2713*	.0294	-.4334*	.1349***	.139*	.55*
	(6.06)	(6.379)	(.276)	(-4.23)	(1.63)	(3.482)	(7.879)
inc_miq	.1401**	.1003**	-.0425	.0165	.1063	.0589	.2633*
	(2.03)	(1.999)	(-.340)	(.12)	(1.14)	(1.26)	(3.201)
inc_umiq	.2044*	.1357*	-.0159	.0262	.1815**	.0712***	.397*
	(2.92)	(2.725)	(-.132)	(.20)	(1.97)	(1.639)	(4.798)
inc_hiq	.3893*	.271*	.0654	-.1554	.2781*	.1658*	.7017*
	(4.94)	(4.712)	(.478)	(-1.09)	(2.61)	(3.230)	(7.618)
ins_prs	.1757*	.1206*	.0137	-.177**	.0708	.065***	.2464*
	(3.54)	(3.457)	(.160)	(-2.15)	(1.03)	(1.919)	(3.913)
fair/poorh	.6689*	.3975*	.6846*	-.2801	.7275*	.624*	.9913*
	(7.54)	(5.896)	(3.896)	(-1.47)	(6.02)	(9.998)	(9.443)
goodh	.476*	.2975*	.2805*	-.1577	.3832*	.3356*	.7301*
	(7.5)	(6.606)	(2.476)	(-1.46)	(4.13)	(7.338)	(8.417)
vgoodh	.2611*	.1612*	.1465	-.1198	.1888**	.1752*	.385*
	(4.46)	(4.091)	(1.470)	(-1.30)	(2.23)	(4.285)	(4.549)
hsiscore	-.574*	-.403*	-.6864*	1.1199*	-.5119*	-.3378*	-.5807*
	(-4.94)	(-4.166)	(-2.875)	(3.08)	(-3.38)	(-4.105)	(-5.236)
chronic	.1574*	.1398*	.0269	-.3894*	.0454**	.0552*	.2203*
	(12.6)	(13.058)	(1.011)	(-7.64)	(2.28)	(5.177)	(17.585)
disability	.0458*	.0389*	.0124	-.1316*	.0208**	.0145*	.0633*
	(8.11)	(7.573)	(1.049)	(-3.52)	(2.64)	(3.159)	(11.893)
alcddly	.017	.0128	-.0027	-.0156	.009	.3628*	.0483***

	(.77)	(.780)	(-.071)	(-.45)	(.28)	(2.252)	(1.690)
pai_mod	-.008 (-.14)	-.02 (-.470)	-.0579 (-.550)	.1861*** (1.70)	.0534 (.63)	.0508 (1.186)	.1757** (2.438)
pai_inc	-.0273 (-.759)	-.029 (-.759)	.0105 (.114)	.1853*** (1.90)	.0865 (1.16)	.048 (1.331)	.0802 (1.228)
smk_dly	-.0961 (-1.48)	-.0752 (-1.594)	.0607 (.533)	.2326** (2.18)	.0911 (.99)	-.0553 (-1.268)	-.1609** (-2.011)
smk_occ	.0358 (.36)	.015 (.200)	.2189 (1.266)	.2434 (1.57)	.3108** (2.14)	.1816* (2.827)	.096 (.886)
hhsmoke	-.0191 (-.32)	-.0113 (-.261)	-.1332 (-1.293)	-.037 (-.38)	-.1249 (-1.50)	-.071*** (-1.758)	-.0169** (-.233)
pr_nfld	-.1907*** (-1.85)	-.1187*** (-1.636)	-.4467** (-2.30)	-.0528 (-.31)	-.428* (-2.93)	-.4367* (-5.942)	-.231*** (-1.798)
pr_pei	.0496 (.51)	.016 (.225)	.0812 (.460)	-.1403 (-.80)	-.0228 (-.17)	-.0268 (-.455)	-.0332 (-.286)
pr_ns	-.1914** (-2.1)	-.1403** (-2.139)	-.1661 (-.942)	.014 (.08)	-.2963** (-2.34)	-.3094* (-4.233)	-.4259* (-3.561)
pr_nb	-.1372 (-1.47)	-.083 (-1.212)	-.1976 (-1.150)	-.0733 (-.44)	-.2432*** (-1.84)	-.277* (-4.125)	-.3366* (-2.891)
pr_que	.2746* (4.77)	.2408* (5.317)	-.4107* (-4.339)	-.6616 (-5.52)	-.2755* (-3.29)	-.185* (-4.562)	.4473* (7.295)
pr_mb	-.1555*** (-1.79)	-.094 (-1.469)	-.3041** (-2.001)	-.0445 (-.28)	-.3253* (-2.68)	-.333* (-5.6)	-.2476* (-2.445)
pr_sask	-.3688* (-3.71)	-.264* (-3.812)	-.2719 (-1.437)	.2716*** (1.73)	-.3506* (-2.56)	-.3984* (-5.448)	-.6495* (-5.012)
pr_ab	-.2882* (-3.56)	-.214* (-3.751)	-.0754 (-.517)	.1888 (1.42)	-.2306** (-2.01)	-.169* (-3.110)	-.5352* (-5.239)
pr_bc	-.3606* (-4.59)	-.223* (-4.048)	-.3074** (-2.049)	.011 (.08)	-.449* (-4.08)	-.4614* (-8.185)	-.7922* (-7.764)
-Log L	9375.1703	9314.975		9298.992		9620.985	
Alpha (α)	3.022224	3.2948* (4.428)		2.0989			
Class Prob						.1102* (27.161)	.8898* (219.412)

Figures in parenthesis are t-statistics.
The symbols *, **, and *** denote 1%, 5%, and 10% significance, respectively.

Table 8
Utilization of Dentist's Services

Variables	NB1	Hurdle Model		ZINB-Probit		Latent Class	
		Stage 1 Probit	Stage 2 ZTNB	Zero Outcome	Non Zero Outcome	Class 1	Class 2
constant	-.6223* (-5.53)	-.3384* (-2.865)	-.2523 (-1.443)	-1.3638* (-4.06)	-.094 (-.74)	1.169* (4.499)	-.9309* (-8.394)
female	.1623* (6.57)	.2301* (7.918)	.0673*** (1.727)	-.1733** (-2.13)	.1368* (5.02)	.1458** (2.213)	.153* (6.154)
married	-.162* (-5.22)	-.1541* (-4.256)	-.177* (-3.552)	.5839* (4.19)	-.0752** (-2.16)	-.2895* (-3.565)	-.177* (-5.702)
wsd	-.1301* (-3.01)	-.149* (-3.018)	-.1644** (-2.255)	.485* (3.13)	-.0713 (-1.46)	-.2641** (-2.131)	-.1372* (-3.182)
age28	.0616** (2.11)	.04 (1.143)	.1044** (2.204)	.2036*** (1.63)	.1008* (3.07)	-.0385 (-.450)	.0888* (3.001)
age38	.0051 (.11)	-.1425* (-2.775)	.292* (3.789)	.7829* (5.32)	.2385* (4.52)	-.008 (-.063)	.0079 (.172)
age48	-.0592 (-.83)	-.1864** (-2.510)	.1691 (1.424)	.836* (4.69)	.1769** (2.18)	-.173 (-.907)	-.0204 (-.296)
yimmig1	-.185 (-1.55)	-.3879* (-3.222)	.174 (.957)	.6826** (1.99)	-.048 (-.36)	-.1099 (-.418)	-.2088** (-1.925)
yimmig2	.0618 (.62)	-.1314 (-1.099)	.3721* (2.824)	.5158 (1.41)	.182 (1.59)	.4563*** (1.804)	.0347 (.380)
yimmig3	.078** (2.08)	.1268* (2.713)	-.025 (-.399)	-.4318* (-2.92)	-.0114 (-.28)	-.0105 (-.105)	.0892** (2.370)
geou	.0767** (2.38)	.0756** (2.164)	.044 (.905)	-.0239 (-.26)	.062*** (1.74)	-.0106 (-.130)	.0754** (2.335)
geom	.1329* (3.25)	.1204** (2.494)	.1832* (2.889)	.003 (.02)	.1654* (3.71)	.1424 (1.391)	.1384** (3.390)
edu_sec	.2016* (5.57)	.2306* (6.225)	-.0774 (-1.506)	-.7523* (-7.42)	-.1237* (-2.95)	-.0313 (-.358)	.2353* (6.584)
edu_grd	.285* (7.66)	.3566* (8.953)	-.0466 (-.867)	-.9223* (-8.15)	-.0515 (-1.21)	-.0217 (-.245)	.3362* (9.136)
inc_miq	.2093* (4.19)	.1753* (3.811)	.0516 (.772)	-.3736* (-3.58)	.04 (.70)	.0946 (.923)	.1994* (4.14)
inc_umiq	.4172* (8.64)	.4233* (9.150)	.037 (.566)	-.7062* (-6.10)	.124** (2.24)	-.007 (-.07)	.4418* (9.567)
inc_hiq	.5564* (10.66)	.666* (12.267)	.1366*** (1.919)	-1.441* (-6.39)	.2381* (4.05)	.1795 (1.544)	.5696* (11.253)
ins_den	.41* (14.56)	.4686* (15.775)	.1519* (3.593)	-.6727* (-7.51)	.2377* (7.4)	.2046** (2.896)	.4503* (16.577)
fair/poorh	-.0989 (-1.60)	-.1823* (-2.749)	.0842 (.900)	.4355* (2.59)	.0322 (.46)	.156 (1.093)	-.1023*** (-1.694)
goodh	-.0181 (-.52)	-.1335* (-3.251)	.1577* (2.984)	.3634* (2.86)	.0583 (1.53)	.1078 (1.227)	-.0318 (-.908)
vgoodh	.0077 (.27)	-.0606*** (-1.721)	.089** (1.960)	.1796 (1.49)	.0274 (.88)	.0377 (.485)	.0095 (.319)
hsiscore	.075 (.81)	.107 (1.112)	-.0415 (-.275)	.066 (.28)	.0623 (.59)	.4964** (2.118)	.1301 (1.425)
chronic	.0258* (2.86)	.0255** (2.382)	.015 (1.066)	-.0323 (-1.18)	.0156 (1.53)	.0302 (1.384)	.0292* (3.207)
disability	.0006 (.12)	-.0013 (-.238)	.0083 (1.112)	.0027 (.20)	.0041 (.77)	.0136 (1.119)	.0023 (.469)

alcddly	-.0294** (-2.20)	-.0386* (-2.586)	-.019 (-.860)	.1101* (2.89)	-.0128 (-.83)	-.0359 (-1.023)	-.0289** (-2.082)
pai_mod	-.046 (-1.42)	-.0934** (-2.351)	.2051 (.040)	.1406 (1.10)	-.0354 (-1.00)	-.7106 (-.798)	-.0445 (-1.35)
pai_inc	-.098* (-3.33)	-.1817* (-5.154)	.0434 (.977)	.3557* (3.24)	-.0257 (-.80)	-.0753 (-.928)	-.1068* (-3.586)
smk_dly	-.1142* (-2.95)	-.1662* (-3.907)	.0344 (.605)	.5293* (4.37)	.024 (.57)	-.1179 (-1.220)	-.1175* (-3.099)
smk_occ	-.0433 (-.76)	-.069 (-1.045)	-.0199 (-.209)	-.0403 (-.16)	-.047 (-.77)	-.1913 (-1.090)	-.004 (-.075)
hhsmoke	-.067*** (-1.90)	-.133* (-3.375)	.0634 (1.226)	.2076*** (1.84)	-.013 (-.34)	.1126 (1.250)	-.066*** (-1.905)
pr_nfld	-.2475* (-3.99)	-.386* (-5.792)	.1562*** (1.785)	.9283* (5.72)	.0575 (.82)	.0237 (.181)	-.3067* (-5.021)
pr_pei	-.0076 (-.13)	-.0074 (-.108)	-.0165 (-.193)	.204 (1.09)	.023 (.37)	-.0572 (-.406)	-.0186 (-.321)
pr_ns	-.1805* (-3.26)	-.251* (-3.977)	-.029 (-.347)	.624* (3.79)	-.04 (-.66)	.0437 (.328)	-.1951* (-3.488)
pr_nb	-.2149* (-3.75)	-.2497* (-3.738)	-.1967** (-2.086)	.8174* (4.88)	-.075 (-1.16)	-.334** (-2.091)	-.2231* (-3.779)
pr_que	-.1816* (-4.82)	-.1907* (-4.393)	-.1977* (-3.430)	.3371* (2.57)	-.1495* (-3.61)	-.1208 (-1.237)	-.2062* (-5.394)
pr_mb	-.0733 (-1.45)	-.113** (-1.937)	-.0012 (-.016)	.1033 (.59)	-.0445 (-.82)	.1178 (.972)	-.1034** (-2.045)
pr_sask	-.3036* (-5.27)	-.3446* (-5.703)	-.141*** (-1.713)	.4472** (2.50)	-.207* (-3.29)	-.0046 (-.033)	-.3442* (-6.013)
pr_ab	-.2587* (-5.68)	-.3685* (-7.081)	-.0099 (-.152)	.6306* (3.85)	-.1137** (-2.28)	.048 (.472)	-.303* (-6.649)
pr_bc	.0003 (.01)	-.0475 (-.912)	.048 (.700)	.0553 (.32)	.004 (.09)	-.0989 (-.913)	-.003 (-.074)
-Log L	13757.248	13502.80		13649.3		13564.92	
Alpha (α)	.6298126	.8655* (11.145)		.3093			
Class Prob						.0474* (12.885)	.9526* (259.14)

Figures in parenthesis are t-statistics.
The symbols *, **, and *** denote 1%, 5%, and 10% significance, respectively.

Table 9
Nights Spent in Hospital

Variables	NB1	Hurdle Model		ZINB-Probit		Latent Class	
		Stage 1 Probit	Stage 2 ZTNB	Zero Outcome	Non Zero Outcome	Class 1	Class 2
constant	-.8438* (-2.95)	-1.3964* (-8.535)	1.0508* (2.619)	1.1523* (6.23)	1.039** (2.92)	2.1131* (30.078)	-2.2342* (-6.571)
female	.2486* (3.06)	.149* (3.196)	-.0377 (-.302)	-.15* (-2.94)	.0339 (.30)	-.1717* (-8.134)	-.5151* (-5.167)
married	.4188* (3.83)	.2293* (4.034)	-.1109 (-.725)	-.265* (-3.96)	-.0685 (-.5)	-.201* (-6.508)	-.1334 (-.867)
wsd	.1454 (1.05)	.083 (1.129)	-.1067 (-.55)	-.1138 (-1.35)	-.103 (-.61)	-.2185* (-6.349)	.0119 (.073)
age28	-.434* (-4.41)	-.2503* (-4.534)	.4338* (2.636)	.366* (5.57)	.4633* (3.18)	.305* (8.620)	-.2913*** (-1.754)
age38	-.0893 (-.70)	-.0818 (-1.124)	1.0421* (5.224)	.2598* (3.11)	1.0726* (6.18)	.701* (19.5)	.3342*** (1.820)
age48	.5078* (3.39)	.244* (2.608)	1.2394* (5.44)	-.0715 (-.69)	1.2655* (6.20)	.8782* (22.268)	1.1004* (5.891)
yimmig1	-.1333 (-.32)	-.0942 (-.422)	-.398 (-.266)	.0176 (.06)	-.436 (-.73)	-.7104* (-2.584)	-25.15 (0)
yimmig2	.029 (.07)	-.0213 (-.097)	.476 (.915)	.121 (.54)	.5798 (1.14)	.2681* (3.71)	-31.07 (0)
yimmig3	-.1822 (-1.36)	-.0554 (-.787)	-.304 (-1.457)	.016 (.19)	-.2916*** (-1.65)	-.2442* (-6.857)	-1.15* (-4.160)
geou	-.0301 (-.33)	-.0036 (-.067)	.0619 (.424)	.0156 (.26)	.0695 (.56)	.1482* (5.406)	.2898** (2.515)
geom	-.3213** (-2.36)	-.1817** (-2.399)	.1518 (.733)	.2231* (2.65)	.1506 (.84)	.1801* (5.270)	-.4401*** (-1.963)
edu_sec	.1341 (1.38)	.0657 (1.183)	.1188 (.813)	-.0604 (-.98)	.0949 (.75)	.0467** (1.964)	.0786 (.765)
edu_grd	.046 (.43)	.0308 (.512)	.0597 (.380)	-.034 (-.50)	.0305 (.22)	.0345 (1.166)	-.0768 (-.582)
inc_miq	-.2206*** (-1.92)	-.136** (-2.099)	-.152 (-.892)	.1214*** (1.65)	-.1786 (-1.20)	.0478 (1.575)	-.0265 (-.224)
inc_umiq	-.105 (-.87)	-.0814 (-1.206)	-.1368 (-.698)	.0618 (.81)	-.171 (-1.07)	-.0246 (-.694)	-.1275 (-.836)
inc_hiq	-.1498 (-1.01)	-.0864 (-1.056)	-.3318 (-1.399)	.0224 (.24)	-.409** (-1.99)	-.2485* (-5.018)	-.1572 (-.714)
ins_hsp	-.1359*** (-1.65)	-.0704 (-1.491)	.1605 (1.238)	.1041** (1.97)	.1662 (1.48)	-.0072 (-.303)	-.4661* (-4.386)
fair/poorh	.5733* (3.61)	.3116* (3.507)	.523** (2.273)	-.2774* (-2.77)	.4917* (2.46)	.534* (12.971)	1.4969* (6.355)
goodh	.3444* (2.84)	.1562** (2.361)	.2754 (1.40)	-.1335*** (-1.79)	.2526 (1.51)	.253* (6.785)	.6545* (2.753)
vgoodh	.0841 (.73)	.0446 (.747)	.13 (.695)	-.0268 (-.39)	.1403 (.90)	.0301 (.793)	.192 (.776)
hsiscore	-.6284* (-3.15)	-.3482* (-2.942)	-.6449* (-2.613)	.3037** (2.27)	-.647* (-2.91)	-.7077* (-19.326)	-1.6353* (-10.218)
chronic	.0927* (3.89)	.0649* (4.564)	.0036 (.099)	-.0714* (-4.38)	.0065 (.20)	-.032* (-5.773)	-.034 (-1.221)
disability	.0803* (8.66)	.0477* (7.846)	.024*** (1.725)	-.0508* (-7.11)	.0218*** (1.91)	.022* (9.282)	.1342* (16.983)
alcddly	-.0543	-.0254	.047	.034	.0453	.1398*	.0799***

	(-1.13)	(-1.134)	(.836)	(1.29)	(.81)	(10.498)	(1.758)
pai_mod	-.111 (-.95)	-.059 (-.932)	.1265 (.723)	.0785 (1.11)	.102 (.64)	.0523 (1.428)	-.3728** (-2.239)
pai_inc	.0325 (.33)	.0126 (.227)	-.0051 (-.033)	-.0249 (-.40)	-.0415 (-.29)	-.0028 (-.088)	-.0038 (-.028)
smk_dly	-.0271 (-.22)	.004 (.055)	-.216 (-1.043)	-.0391 (-.51)	-.222 (-1.37)	.1342* (3.579)	.3844*** (1.672)
smk_occ	-.0312 (-.16)	.0074 (.071)	-.1487 (-.389)	-.0153 (-.12)	-.0885 (-.33)	-.2302** (-2.166)	-1.455 (-1.592)
hhsmoke	-.0757 (-.66)	-.051 (-.837)	.1777 (.953)	.085 (1.21)	.1741 (1.18)	-.0895* (-2.767)	-.457** (-1.986)
pr_nfld	-.1198 (-.62)	-.0857 (-.837)	.1112 (.407)	.109 (.95)	.1038 (.41)	-.0814 (-1.557)	-.7764** (-2.080)
pr_pei	.2943*** (1.75)	.1575*** (1.644)	.2672 (1.181)	-.136 (-1.27)	.2502 (1.13)	.494* (12.356)	.458** (2.544)
pr_ns	-.1811 (-1.01)	-.116 (-1.202)	.2295 (.830)	.1641 (1.54)	.2468 (1.07)	.0821*** (1.702)	-.0637 -.391
pr_nb	.5256* (3.57)	.2725* (3.082)	-.109 (-.475)	-.3284* (-3.19)	-.1197 (-.61)	-.096** (-2.198)	.4196** (2.276)
pr_que	.275** (2.30)	.145** (2.211)	.1166 (.682)	-.1502** (-2.05)	.0942 (.61)	.135* (4.843)	.2364 (1.555)
pr_mb	.2088 (1.38)	.075 (.857)	.1063 (.467)	-.0632 (-.65)	.1102 (.55)	.1927* (4.697)	.543* (3.466)
pr_sask	-.0698 (-.41)	-.0488 (-.521)	-.046 (-.151)	.059 (.56)	.0114 (.05)	-.2067* (-3.734)	-.935* (-3.025)
pr_ab	-.08 (-.53)	-.0512 (-.626)	-.4795 (-1.987)	-.0314 (-.32)	-.4741** (-2.36)	-.4288* (-8.131)	-.1403 (-.708)
pr_bc	-.0683 (-.46)	-.0103 (-.132)	-.341 (-1.56)	-.0486 (-.52)	-.3531*** (-1.80)	-.4368* (-8.694)	-.3897*** (-1.844)
-Log L	4310.0066	4213.389		4214.378		4778.520	
Alpha (α)	14.40082	1.326* (6.599)		1.355			
Class Prob						.0541* (22.593)	.9459* (395.327)

Figures in parenthesis are t-statistics.
The symbols *, **, and *** denote 1%, 5%, and 10% significance, respectively.

1 Introduction

Understanding the underlying process of the demand for health care is quintessential for a better assessment of the role of public intervention in the health sector . This issue is gaining momentum in recent years, especially in developing countries of the world. Governments, public policy makers, economists, and people around the world are debating who should pay for what; and how best to organize and deliver health services so as to allocate scarce resources efficiently and work towards a healthier society. Moreover, from a development perspective, design of health policy is of utmost importance because health status of the population in developing countries generally, is far below that of the developed countries in general.

It is well known that public spending on health care remains one of the most uncontroversial roles of the government in developing countries of the world. There are several arguments in favour of this. First, for certain kinds of services there are significant externalities so that the competitive market will not provide an optimal amount. For instance, market mechanism may fail in the prevention and treatment of infectious diseases (like tuberculosis, malaria, plague, small pox, HIV, etc.), and in the provision of regular public health measures such as safe drinking water, public toilet facilities, safe waste disposal, vaccination programs, health awareness campaigns. Second, public spending on health care is advocated on the grounds of equity, especially when there exist large disparities in per capita income. Since low income groups are at high risk of illness in general, public financing of health care is an effective way of redistributing income in favour of the poor. Third, in most developing countries, the corrupt bureaucratic practice coupled with the inefficient administration and a weak tax system limits the ability of any program that is geared up for direct cash provision. So, government provision of services would likely be favoured as an efficient mechanism design policy.

In many developing countries there is extensive public support for hospitals, medical education, drugs, etc. Most of the health services are typically provided at little or no monetary prices. However, what is not so clear is whether governments spend appropriately in order to raise access to and use of health care regardless of ability to pay. In this paper, we ask the following set of interrelated questions. What are the determinants of demand? How important are price, income, quality, and access in the choice of a health care provider? How do rich and poor individuals make decisions about their treatment in response to price? What are the implications for equitable access to health care and health status of the people across income groups?

The paper is organized as follows. In section 2, I present a brief overview of the Indian health care system. Section 3 discusses the demand for health care and reviews the related literature. Section 4 deals with the econometric methods. The data source and variable construction is discussed in section 5. In section 6, the estimated results are reported and discussed. Finally, section 7 is reserved for discussions of conclusions, future research works, and limitations of this study.

2 Indian Health Care System

2.1 The Context

The Indian health care system has brought the 'health status' of Indian people far from where what it inherited at the time of independence. The male life expectancy at birth has increased from 27 years at the time of independence to 62 by the end of 1996; that of females has been increased from 27 to 63 during the same period. The infant mortality rate declined from 162 per 1000 population to 72 by the end of 1996. Consequently, the death rate has decreased from 14 to 8.9 per 1000 population. Burdens of some communicable diseases have declined significantly and some of them have been eradicated. All these achievements and others that are shown in Table 1 were possible because of a tremendous increase in health resources to the population since independence coupled with better health seeking behaviour by the people. In addition, rapid improvements in health indicators may also be attributed to an improvement in the population general well being. For instance, doctors per 10,000 population has increased from 1.6 during 1947 to 4.2 during 1988; midwives per 10,000 population has increased from 0.2 to 1.6; nurses have been increased from 0.2 to 3, and the supply of other health visitors/workers has increased.

Better immunization coverage, increased access to safe drinking water, and better supply of drugs are other contributory factors that prevented ill-health and disease. Increasing access to education, awareness, information, and other demand factors led to higher use of both preventive and curative health care. However, there is a great potential to improve the health status of people. For instance, in many developed countries, the life expectancy at birth is close to 80, the infant mortality rate is below 20, and mortality due to communicable diseases is almost negligible.

With an estimated population of more than 1 billion people, India remains first in the world in terms of the number added to its population each year - about 16 million. Malnutrition also poses a continuing constraint to India's development. More than half of India's children under four and 30 percent of newborns are significantly underweight. Despite some improvement, India's women remain significantly more malnourished than men, and 60 percent of Indian women are anemic. Bias against women and girls is reflected in the demographic ratio of 929 females for every 1,000 males. Unlike most countries, more women than men die before the age of 35 in India. Although declining, largely preventable diseases such as tuberculosis, cataract blindness, and malaria continue to account for 50 percent of reported illness, and around 470 deaths per100,000. HIV/AIDS is a newly emerging threat to India's public health. According to the most recent report from the National AIDS Control Organization, India has more than 4 million HIV infected people.

As far as the institutional arrangements in the delivery of health care are concerned, there are many providers offering varying degrees of services in rural and urban areas. Among others, Primary Health Care Centre (PHC), public hospitals, public dispensaries, private & voluntary institutions, and private doctors provide varieties of health services in India. In addition to public financing, private and voluntary organizations and individual households contribute substantially to the financing of health care.

2.2 The Pattern of Illness

The National Council for Applied Economic Research (NCAER) has worked out the prevalent rate of illness (defined as the illness for the one-month reference period), for major States in India. On the basis of a Household survey, NCAER pointed out that the prevalence rate of illness per 1000 population is relatively higher in rural areas compared to urban areas: 107 versus 103 (NCAER [40]). Prevalence rate of male illness is 106 and 98 in rural and urban areas respectively. Prevalence rate of female illness is 108 in rural areas and 99 in urban areas. These numbers clearly show that the burden of illness is relatively higher in rural areas where less attention is paid. A study made by Krishnan [32], on the basis of NSS 42nd survey, shows that except for Kerala rural patients pay more for health care and bear a higher burden of treatment, reflecting rural-urban bias in health facilities in all states.

On the basis of the NCAER survey, 62% of the hospitalization care in urban areas, and 60% in rural areas are met by the public sector and the rest is met by the private sector. The NSS 42nd survey also shows the same pattern, around 60% in both rural and urban areas. However, as far as the inter-state variation is concerned, these two surveys show striking differences. The NSS 42nd survey shows that the states with better health care infrastructure have a lower percentage treated in public hospitals; this is particularly true for Kerala and Maharashtra. In states like Orissa and Uttar Pradesh, where the health care infrastructure is not so developed, the percentage of inpatients treated in government hospitals is around 80%. Approximately 40% of the demand for outpatient care (i.e., non-hospitalized care), is met by the public sector in rural areas, and the corresponding figure in urban areas is 34%.

There are many reasons to explain this two-tier structure for outpatient care. First, for outpatient care the opportunity cost of time (income and work foregone) may be higher in public sector. Second, the nature of medical care might not be of expected quality in the subjective evaluation of patients. Third, some people might be so rich that they really don't care for any public facility and always consult a specialist in the private sector. Fourth, in some instances, people might be using traditional medicines and consulting traditional physicians or resorting to self medication for non-serious illnesses.

Vishwanathan and Rhode [48] found that 65% of the diarrhea patients chose private medical practitioners. Yesudian [49], in a survey of the utilization of the medical facilities by the slum dwellers in Bombay, found that people use private facilities more frequently for short term and minor ailments. However, for acute illnesses requiring hospitalization, they used public facilities. Similar results are also found at two locations in the Ganjam district of Orissa (Sarma [43]).

Regarding costs and quality, Uplekar [45] found that private physicians serving the urban poor in the slums of Bombay do not consider standard, recommended drugs in treating pulmonary tuberculosis. Upelkar [46] found that the private physicians had inadequate knowledge about treatment of leprosy. Greenhalgh [17], in a survey of 2,400 patients treated by public and private providers, observed that private doctors prescribe more drugs. He found that some specialized drugs are prescribed which are often used inappropriately. He also observed that 64 percent of the patients bought medicines over the counter without having a prescription. Duggal and Amin [14], in their survey in the village Jalgaon, observed that physicians in the private sector use more injections and medicines than physicians in the public sector.

2.3 A Profile of the Incidence of Non-Hospitalized Ailments and Treatment Decisions

The present sub-section is based on the survey information on morbidity and health care conducted in the 52nd round of the NSSO during July 1995 - June 1996. Table 2 in the Appendix gives the sample figures and the corresponding population estimates on prevalence of self-reported morbidity, the number of persons reporting an ailment (chronic or acute) during 15 days and the corresponding numbers per 1000 persons, for males and females in both rural and urban areas.[1] Although the findings indicate that there is hardly any difference in the morbidity rates between rural and urban areas, there seems to be perceptible gender differences in each of rural and urban areas. It is also evident from Table 2 that gender specific rates for acute ailments were about almost three times higher than that of the chronic ailments in both rural and urban areas.

According to the NSSO [41] report, the age-specific morbidity rates for acute ailments shows a U-shaped relationship, whereas for chronic ailments it is positively-sloped; as expected both ailments are found to be much higher

[1]It is to be noted that the normal pregnancy and child birth related treatments were not treated as ailments in the survey. However, pregnancy and childbirth complications were treated as ailments.

among those aged 60 years and above. The NSSO [41] report also indicates that there is a positive relationship between monthly per capita consumption expenditure and morbidity rates in both rural and urban areas. If the consumption expenditure is considered to be a proxy for income (or as the report treats this as the level of living of the households), then the results suggest that the level of morbidity tends to rise with income or the level of living. This may be either because the poor are healthy (quite unlikely), or that the reporting of morbidity rises with income (quite likely).

The proportion of ailing persons treated during the reference period is found to be higher in urban areas (91%) than in rural areas (83%), and seems to have no gender differences. Among those who didn't seek treatment, about 26% in the lowest expenditure group to about 10% in top expenditure group were found in urban sample; whereas the corresponding figures in urban areas were 19% to 9%, respectively. Regarding the reasons associated with no treatments, 52% and 60% of the untreated ailments were reported as non-serious in rural and urban areas, respectively. The next most important explanation for no treatment was financial problems, accounting for 24% and 21% in rural and urban areas, respectively. The other explanations for no treatments were no medical facility, long waiting, lack of faith, and others.

Table 3 in the Appendix gives the distribution of outpatient treatments by source of treatment in rural areas. It is evident from Table 3 that private doctors were the most dominant source of outpatient treatment in both rural and urban areas. Almost 80% of the outpatient treatments were in the private sector, comprising of private doctors, private hospitals, nursing homes, charitable institutions, etc. This indicates that there is a sharp rise in the share of private sector for outpatient treatments over time, as evident from the findings of 42nd round of NSSO survey. The share of the public sector for outpatient treatments on the other hand constitute about 20%. The public sector includes government hospitals, clinics, dispensaries, PHCs, CHCs, and the Central and State government aided ESI facilities.

3 Demand for Health Care

3.1 Introduction

The notion of demand for health care is closely related to the health seeking behaviour of individuals in any society. Before a person consume any medical care, either from private sources or from government sources or self care, she/he must perceive the need for it and then demand it. Need for health care could be either ‘self-perceived’, or ‘observed’. For instance, pain, headache, hygienic behaviour, psychosomatic problems, etc. that are internal to the individual herself/himself are examples of self-perceived need; while observed need is any thing that can be observed and assessed by a trained individual, which might or might not have been perceived by the concerned individual. In any case, when there arises a need for health care, individuals decide whether to visit a doctor and where to visit. The process of making such decisions may be complicated because of little information or too much information from friends, relatives, neighbours, physicians, and advertisements about the potential costs, risks, benefits, and opportunity cost of foregoing consumption of non-medical commodities.

3.2 The Literature

In the health economics literature, two alternative approaches are used to model this complicated decision making processes regarding health care utilization. One approach to model health care choices is to use an intertemporal model of consumption decisions and treat health as a stock variable (Grossman [21]). In this approach, health care is demanded to the extent that it improves the stock of health and increases productivity. The second approach is to treat health care as only one of the several commodities over which economic agents have well defined preferences, Phelps [42]. We follow the second approach to model preferences for health where health care would be demanded as an input into the production of health. We then analyze the effect of price, income, and health status on the demand for health care. An alternative way to model preferences for health care is to use a state dependent approach, Zweifal and Bryer [50]. However, our data set do not permit us to pursue a state dependent approach.

The utility function is defined as $U(c, h)$, where c is a generic consumption good other than health care and h is the level of health. Assume that the utility function is well defined, that is $U_c > 0, U_h > 0, U_{cc} < 0, U_{hh} < 0$. Health care is demanded only to the extent that it improves the underlying health of the individual, the effectiveness of which is determined by a host of factors including the health status. If θ units of health care are needed in order to produce an additional unit health, then the effect of sickness implies that the value of θ increases, and given income and prices, the equilibrium pair $\{h, c\}$ shrinks. Following Jack [29], if the price elasticity of demand for health lies between zero and one, the one-to-one relationship between $\{c, h\}$ space and $\{c, s\}$ space can be represented through Figure 1 in the Appendix, where s represents health care. In Figure 1(a), when the person is well, the equilibrium pair is represented by $\{c_1, h_1\}$, when ill the new equilibrium pair is $\{c_2, h_2\}$. If the person would like to choose c_1 levels of consumption in the event of illness, then the level of health will be reduced to h_3. At point $\{c_1, h_3\}$, the indifference curve must be less steep than through $\{c_2, h_2\}$, which is less steep than through $\{c_1, h_1\}$. The corresponding equilibrium pair of consumption and health care when well is represented as $\{c_1, s_1\}$ in Figure 1(b), the bold indifference curve. The equilibrium pair $\{c_2, h_2\}$ in $\{c, h\}$ space corresponds to the pair $\{c_2, s_2\}$ in $\{c, s\}$ space. If the price of consumption and health are normalized to unity and the price elasticity of demand for health is between 0 and 1 then the pair $\{c_1, h_3\}$ in $\{c, h\}$ corresponds to an indifference curve that cuts the bold indifference curve at $\{c_1, s_1\}$. On the other hand, if the price elasticity of demand for health is greater than 1, then the indifference curve must cut the bold indifference curve from above in the $\{c, s\}$ space. The implict assumption is that the effect of illness is to increase the price per unit of health. However, incidence of an illness might affect an individual's earning thereby leading to potential income effects.[2] The crux of the argument however is that price, income, and health status are likely to

[2]See Jack [29] for an excellent exposition of a number of related conceptual issues in modelling the demand for health care in developing countries.

affect the demand for health care and their magnitudes are unclear *a priori* but most important in designing health policy in developing countries.

The literature on the demand for health care is not confined to quantities of health care, but most importantly choice of provider type. The existence of more than one type of provider means a somewhat different analytical framework is needed to estimate demand functions. In many instances, individuals are able to choose from a set of alternative providers, where each provider-choice leads to a potential improvement in expected health for a price. The price of an alternative in turn may include both monetary (medical and non-medical expenses including loss of income) and non-monetary costs. Taking into account all these information, income, and health status, the rational decision maker chooses the alternative that yields the highest expected utility. More precisely, the expected utility conditional on choosing an alternative, say j, can be written as $U(c_j, h_j)\forall j$, where h_j is the expected improvement in health after receiving care from provider j and c_j is the consumption net of costs of care from provider j. We shall return to functional form, identification, and estimation issues in the next section.

There are a number of studies on the demand for medical care in developing countries. Some studies found that prices are not important determinants of demand for medical care (Akin et al. [2], Akin et al. [3], Schwartz et al. [44], Birdshall and Chuhan [5], Heller[23]); while some other studies found that prices are indeed important determinants of demand for medical care (Mwabu [37], Mwabu et al. [38], Mwabu et al.[39], Dor et al. [10], Gertler et al. [18], Gertler and van der Gaag [19], Bolduc et al. [7], Dow [12], Dow [13]). All these studies employ discrete choice models to analyze the choice of health care provider. However, the methods and results on the price and income elasticities are confounding across studies thereby making general policy implications difficult and sometimes even inconsistent.[3] Many of the studies contradict the findings from the developed countries where price elasticities range from -0.2 to -2.1. These conflicting findings may seem to be paradoxical because one might expect price elasticities to be higher in developing countries due to low income and high uninsured population (Gertler and van der Gaag [19]). On the contrary, price elasticities may not be higher because price per unit of care is much lower in developing countries; and more importantly the health seeking behaviour of the people in developing countries might not correspond to that of the behaviour patterns of developed countries.

[3] See Gretler and van der Gaag [19], Jimenez [30], and Gretler and Hammer [20].

3.3 Limitations of the Previous Literature

This apparent conflicting findings may be explained through the following limitations of the previous literature. The previous empirical literature on the demand for health care does suffer from a number of issues. First, the treatment of the price of medical care appears to be completely inadequate. Some studies have used standard fee schedules as reported by the provider, some studies used expenditures per medical visit as the relevant price, and other studies used results from a hedonic price equation. However, all these methods can cause misleading results. This is because standard official fees may not reflect the true price of care for several reasons. The expected amount spent by a person for a specific illness may depend not only on the standard fees but also the type of treatment, quality of treatment, individual idiosyncratic elements, and other non-medical expenses chosen by the patient. It is not unusual in the health care market that a range of treatments are provided with varying degrees of price and quality for the same ailment. Thus, in order to compute the true prices for each provider choice faced by an individual, one must account for all these elements including uncertainty.

Second, all the models of discrete choice of health care demand to date employed a choice set that is fixed across individuals regardless of the true choice generating processes. This is a serious limitation because not all provider types are capable of handling all types of ailments, not all provider types are accessible in every conceivable geographical location, and not all provider types are even affordable by everyone. In other words, a choice generating process has to be defined for each decision maker confronting a health care provider choice setting situation.

Third, most of the studies used a nested multinomial logit framework to model health care provider choice, but alternative scaling and forms of degenerate and partially degenerate models can obscure comparability; certain forms are even inconsistent with the utility maximization hypothesis; and invariance cannot be achieved unless certain restrictions are imposed. In this paper, we address these specific concerns in the context of outpatient health care demand in rural India.

The paper is intended to contribute to the literature on health care demand in several ways. Use of the discrete choice model to explain the demand for outpatient health care on the basis of NSS data is entirely new. Most of the literature on multinomial choice model of health care demand has been restricted to a situation in which the choice set is fixed across individuals. In the context of a country like India, the true choice generating process may vary across individuals by geographical location, nature of illness, and affordability. Thus, we generated a variable choice set to reflect the true choice generating process. In most of the survey data, we observe only the price of care for the chosen alternatives, and we need the price information about unchosen alternatives to model the demand for health care. In this paper, the price of care for an unchosen alternative in the choice set for an individual is imputed as a mean of a random sample with replacement from the specific provider, province, illness, and income group. The random sample is drawn for each individual's unchosen alternatives to reflect upon individual heterogeneity and uncertainty about the expected price of care faced by an individual confronted with a choice setting situation. This paper also addresses recent econometric ideas relating to identification, scaling, invariance, and consistency with the utility maximization hypothesis that underlies the basis of modelling health care demand.[4]

4 The Model

Let C be the universal choice set that includes all possible choices for some population. However, for a particular individual $n \in N, N = \{1, 2, ..., N\}$,

[4]See Hunt [28] and Hensher and Greene [27].

the relevant choice set is $C_n \subseteq C$. This is because the nature of illness may be such that some providers will not be chosen by some individuals or some providers may not be accessible to some individuals or a particular choice set may not be feasible in terms of affordability. Let J be the number of elements in C, and $J_n \leq J$ be the number of elements in C_n. When individuals need medical attention, they are faced with the above alternatives, and a choice must be made. An individual faced with a set of feasible alternatives chooses the one that yields the highest utility. The observed attributes of alternative $j \in J_n$ faced by the patient $n \in N$ as the vector $\mathbf{z}_{jn} \forall j \in J_n$. Different patients might possibly make different choices when confronted with the same alternatives, because the subjective valuation that they place on each possible alternative is different. The differences in the valuation of each alternative may depend on the specific characteristics of the alternative, or specific characteristics of the decision maker; some of those are observed and others are unobserved. Let the observed characteristics of the patient n as the vector $\mathbf{x}_n$. The probability that patient n chooses alternative $j \in J_n$ then depends on the observed attributes of alternative j and the observed characteristics of the decision maker. Let us denote this probability as π_{jn}.

Modelling discrete choice situation essentially involves specifying π_{jn} as a parametric function of the general form: $\pi_{jn} = f(\mathbf{z}_{jn}, \mathbf{x}_n, \boldsymbol{\beta})$. Let U_{jn} be the utility of choosing alternative j by patient n, which depends on the observed attributes of alternative $j, \mathbf{z}_{jn}$; the observed characteristics of the patient, $\mathbf{x}_n$; and some unobserved characteristics that are not known. If everything is known, then the deterministic utility function can be specified as

$$U_{jn} = U\left(\mathbf{z}_{jn}^*, \mathbf{x}_n^*\right) \forall j \in J_n. \tag{1}$$

Where, $\mathbf{z}_{jn}^*$ is the all relevant attributes of alternative j faced by the patient n and $\mathbf{x}_n^*$ is all relevant characteristics of the patient n. The patient n chooses the alternative from which she/he derives the maximum utility. Alternative $j \succcurlyeq i$ iff $U_{jn} \geq U_{in} \forall i \in J_n, i \neq j$. In this deterministic setting, the probability that the patient n chooses alternative j is either one or zero depending on whether the alternative j gives the maximum utility or not. The idea is that if we know all the relevant factors and the preferences of the patient we could effectively predict the patient's choice of a provider. But in practice, we simply do not know all the relevant factors and the form of the exact utility function. What we observe is only $\mathbf{z}_{jn} \subset \mathbf{z}_{jn}^*$ and $\mathbf{x}_n \subset \mathbf{x}_n^*$, the relevant sub vectors of the alternative-specific and individual-specific variables, respectively. We can therefore bifurcate the utility function into two sub-functions, one that is known up to a vector of parameters $\boldsymbol{\beta}$ to be estimated denoted as $V(\mathbf{z}_{jn}, \mathrm{x}_n, \boldsymbol{\beta})$, and the other that represents all factors and aspects of utility and alternative characteristics that are unknown, denoted as ε_{jn}. We can now specify the patient's utility function as

$$V_{jn} = V(\mathbf{z}_{jn}, \mathrm{x}_n, \boldsymbol{\beta}) + \varepsilon_{jn}. \tag{2}$$

We assume that the deterministic part of the utility function is known and we want to predict patient's choice based on this limited information.

If $U_{jn} \geq U_{in}\ \forall i, j \in J_n$ then alternative j is chosen, otherwise, some other alternative is chosen. Thus, the probability of choosing alternative $j \in J_n$ by individual $n \in N$ is

$$\pi_{jn} = \mathrm{pr}\left[V_{jn} + \varepsilon_{jn} \geq \max_{\substack{i \in J_n \\ i \neq j}} \{V_{in} + \varepsilon_{in}\}\right]. \tag{3}$$

Let the deterministic utility conditional on receiving care from alternative $j \in J_n$ be given by

$$V_{jn} = V(h_{jn}, c_{jn}). \tag{4}$$

Where, h_{jn} is the expected improvement in health after receiving care from alternative j and c_{jn} is consumption net of costs of health care. Following Gertler et al [18], let h_{0n} be the expected health from a reference alternative (for example, self care). Therefore, the change in expected improvement in health from choosing alternative j rather than the reference alternative is $h_{jn} - h_{0n}$. If $h_{jn} - h_{0n}$ is positive then alternative j is supposed to have a positive impact on health of patient n. Let us denote this change in expected improvement in health from choosing alternative j is E_{jn}, i.e., expected effectiveness or quality measure of alternative j. Therefore, the expected health production function is given by

$$h_{jn} = E_{jn} + h_{0n}. \tag{5}$$

In fact, E_{jn} depends on $\mathbf{x}_{jn}$, which includes educational status, health status, severity of illness, and other patient characteristics, and $\overline{\mathbf{z}}_{jn}$. That is,

$$E_{jn} = E\left(\mathbf{x}_{jn}, \overline{\mathbf{z}}_{jn}\right). \tag{6}$$

Thus, the conditional utility function can be specified as

$$U_{jn} = V\left(E\left(\mathbf{x}_{jn}, \overline{\mathbf{z}}_{jn}\right) + h_{0n}\right) + \varepsilon_{jn}. \tag{7}$$

The unconditional utility maximization problem for patient n is thus specified as $U_n^* = \max_{i \in C_n} \{U_{in}\}$, where U_n^* is the highest utility that patient n can obtain. The solution to this problem yields a probability choice systems, i.e., a system of demand functions for alternatives. The functional forms of the demand functions depends on the functional form of the utility function and the distribution assumption of the stochastic term.

4.1 Econometric Specifications

It is customary to begin with a linear functional form:

$$U_{jn} = \alpha_1 c_{jn} + \alpha_2 h_{jn} + \varepsilon_{jn}$$

The individual faces a budget constraint such that consumption plus the price of health care must be less than or equal to income, which implies that $c_{jn} = Y - P_{j_n}$. Choice is also constrained by the health production function which is ideally a function of both quality and a set of individual characteristics. So, the underlying indirect utility function can be written as

$$V_j = \beta_{0j} + \beta_{1j} Y + \beta_{2j} P_j + \beta_{3j} \mathbf{X}_j + \varepsilon_j, \tag{8}$$

where $\beta_{1j} = \beta_{1k}$, and $\beta_{2j} = \beta_{2k} \forall j, k \in C$. Since income does not vary across choices, $\beta_1 Y$ can be dropped from the estimation or we need to impose the restriction that $\beta_1 = -\beta_2$. However, these implict restrictions are often violated in the empirical studies.[5] Gertler et al. [18] and Gertler and van der Gaag [19] argue that if $\alpha_{1j} \neq \alpha_{1k}$ then two alternatives produce the same amount of health improvement for the same price but yields two different levels of utility. If this is the case, preferences are not well defined and stable utility functions do not exist. The functional form we used in this study is the parsimonious approach of Gertler et al. [18], where prices and income is quadratic in the logs of net income.

$$V_j = \beta_{0j} + \beta_1 \ln(Y - P_j) + \beta_2 \left[\ln(Y - P_j)\right]^2 + \beta_{3j} \mathbf{X}_j + \varepsilon_j. \tag{9}$$

[5]See for instance, Akin et al. [2], [3], Birdsall and Chuhan [5], Dor and van der Gaag [11], Mwabu [37], Mwabu et al. [38].

This functional form relaxes the restriction that income has no effect on provider choice, any assumption about the marginal rate of substitution, and inconsistency with the axioms of utility maximization.[6]

4.2 Distributional Assumptions

Estimation of a choice model depends on the underlying distribution assumptions about the stochastic term. If each $\varepsilon_{jn}, \forall j \in J_n$, is distributed independently with an extreme value distribution then the probability that patient $n \in N$ will choose $j \in J_n$ is $\frac{e^{V_{jn}}}{\sum_{k \in J_n} e^{V_{kn}}}$. This is known as the multinomial logit (MNL) specification.[7] One of the underlying assumptions of the multinomial logit model is that the ratio of the probabilities of the two alternatives, j and k depends only on alternatives j and k, and not on the presence of any other alternatives. This is known as the independence of irrelevance alternatives (IIA) property. If IIA property is false, then the model is misspecified, thereby producing misleading estimates and false conclusions.

[6]However, we retain the assumption that the marginal utility of income to be the same across alternatives because we impose the restriction that β_1 and β_2 to be equal across alternatives. There is a debate about whether constraining alternative specific coefficients to be equal are consistent with utility maximization (see Dow, [13]), but radically different marginal utilities across alternatives may be implausible.

[7]If $V_{jn}(.)$ consists of both alternative-specific and individual-specific variables, then this is also known as mixed-logit model in the literature.

The core assumption that we need in order to derive the multinomial logit model is that the disturbances are independent. If a pair of disturbances are not independent, then we have a severe problem. It is argued that, "any model other than multinomial logit model might produce different numerical results, but any model based on the assumption that all the disturbances are independent are subject to the victim of IIA assumption" (Ben-Akiva and Lerman [4]).

The independence assumption implies that the unobserved component $\varepsilon_{jn}, \forall j \in J_n$ and any $\varepsilon_{in}; \forall i \in J_n, i \neq j$ are assumed to have the same distribution, with the same mean and variance, and they are uncorrelated with each other. That the random variables are uncorrected with each other means that any factor that we do not observe and affects the utility of alternative j does not affect the utility of every other alternative $i \in J_n, i \neq j$. The assumption that the random variables have the same variance means that the unobservable variables that affect the utility of alternative j has the same variation as the different unobserved factors that affect the utility of alternative i (because of the zero correlation). We must therefore search for alternative assumptions that relax this assumption. Two lines of research have been developed to incorporate cross-correlations. One approach follows within the tradition of a closed form analysis while the other approach relies on probability simulation methods to get rid of the curse of dimensionality problems arising in the multinomial probit (MNP) model. Though MNP is a useful candidate, it is computationally intensive for problems with more than a few alternatives because of the evaluation of multiple integrals. Due to severe identification problems, MNP is rarely used beyond three or four alternatives. McFadden [34] introduced the nested multinomial logit model (NMNL) as a compromise between functional flexibility and computational feasibility. There is therefore attraction towards NMNL models because of its ability to incorporate differential degrees of interdependence between subsets of alternatives within a choice set while maintaining IIA assumptions within each subset.

However, two forms of NMNL derivations appear in the literature. One derivation is known as non normalized nested logit (NNNL) model, which is a generalization of MNL model and is inconsistent with utility maximization; and the other is utility maximizing nested logit (UMNL) model, which is derived from the generalized extreme value (GEV) distribution.[8] Here, we shall present the UMNL specification of health provider choice and compare

[8]For details about the derivation and comparison, see Hunt [28] and Hensher and Greene [27].

our results with the other forms when there are partial degenerate branches. A case of complete degeneracy (i.e., MNL) and two level partial degeneracy cases of health care provider choice are presented in Figure 2(a), Figure 2(b), and Figure 2(c) in the Appendix, respectively. The idea behind a nested logit specification is that J_n alternatives are partitioned into subsets of alternatives such that the ratio of probabilities for any two alternatives that are in the same subset is independent of the existence of other alternatives. The assumption of nonindependence of the unobservable components in the nested logit model underlies the rational for a nesting structure. For instance, in Figure 2(b), the choice structure consists of levels **1** and **2**, and one degenerate branch labeled as **11**. For a person $n \in N$, utility associated with the level **2** is $U_{n2}+\nu_{n2}$, and the utility associated with the lower level **1** conditional on upper level **2** is $U_{n21} = (U_{n2} + \nu_{n2})+(U_{n21} + \nu_{n21})$. Similarly the utility associated with choice **22** is $U_{n22} = (U_{n2} + \nu_{n2})+(U_{n22} + \nu_{n22})$. So each choice in a non-degenerate branch shares a common component and a own specific utility component. The nesting structure assumes that the unobserved components at any level and across levels are independently distributed. However, a positive correlation arises among unobservable components that share the same upper level. For instance, $Cov\,(\varepsilon_{n21}, \varepsilon_{n22}) = E[(\nu_{n2} + \nu_{n21})(\nu_{n2} + \nu_{n22})] =$

$E(\nu_{n2}\nu_{n2}) + E(\nu_{n2}\nu_{n22}) + E(\nu_{n21}\nu_{n2}) + E(\nu_{n21}\nu_{n22}) = Var(\nu_{n2})$. In contrast to this, there is no shared unobserved common component in Figure 2(a).

For the sake of exposition, assume that there are no degenerate branches, index m and k be the upper level choices (i.e., $C_n = \{C_m|C_k\}$), j and i be the lower level choices, and drop individual index. The joint probability of choosing an alternative j in branch m is: $\pi(m,j) = \pi(m) \cdot \pi(j|m)$. The probability of choosing an upper level branch is: $\pi(m) = \Pr\left[(U_m + \nu_m) + \max_{j \in C_m}(U_{mj} + \nu_{mj})\right] \geq \Pr\left[(U_k + \nu_k) + \max_{i \in C_k}(U_{ki} + \nu_{ki})\right]$, $\forall m \neq k$. Let $V_m^* = \max_{j \in C_m}(U_{mj} + \nu_{mj})$ and $V_k^* = \max_{i \in C_k}(U_{ki} + \nu_{ki})$. Assuming that $\nu_{mj} \sim IID$ Gumbel $(0, \mu^m)$ and V_m^* follows IID Gumbel with location parameter $\left(\frac{1}{\mu^m}\right) \ln \sum_{j \in C_m} \exp(\mu^m U_{mj})$ $\equiv C_m^*$ and scale parameter μ^m for branch m. Analogous location and scaling parameter can be written for branch k, let the corresponding scaling parameter be μ^k. Thus, $V_m^* = C_m^* + \nu_m^*$ and $V_k^* = C_k^* + \nu_k^*$; $\nu_m^* \sim IID$ Gumbel $(0, \mu^m)$, and $\nu_k^* \sim IID$ Gumbel$\left(0, \mu^k\right)$. This implies that $\pi(m) = \Pr\left[(\nu_k + \nu_k^*) - (\nu_m + \nu_m^*)\right] \leq \Pr\left[(U_m + C_m^*) - (U_k + C_k^*)\right], \forall m \neq k; (\nu_m + \nu_m^*) \sim IID$ Gumbel $(0, \lambda^m)$ and $\nu_k^* \sim IID$ Gumbel $\left(0, \lambda^k\right)$. Since λ represents scaling of the composite unobservables, which includes variances from both the upper and lower level choices and their variances cannot be less than or equal to lower level variances. This means that, $\lambda^m < \mu^m$ and $\lambda^k < \mu^k$.

Assuming common scale parameters across branches, i.e., $\mu^m = \mu^k = \mu$ and $\lambda^m = \lambda^k = \lambda$ and since difference between two Gumbel variates $[(\nu_k + \nu_k^*) - (\nu_m + \nu_m^*)]$ follows a logistic distribution, the cumulative density function yields the following upper bound: $\pi(m) = \exp\left[\lambda\left(U_m + C_m^*\right)\right] / \sum_{\forall k} \exp \lambda \left[U_k + C_k^*\right]$. The inclusive values (IV) are: $IV_m = \ln \sum_{j \in C_m} \exp\left(\mu U_{mj}\right)$ and $IV_k = \ln \sum_{i \in C_k} \exp\left(\mu U_{ki}\right)$. Thus,

$$\pi\left(m\right) = \exp\left[\lambda U_m + \left(\frac{\lambda}{\mu}\right) IV_m\right] / \sum_{\forall k} \exp\left[\lambda U_k + \left(\frac{\lambda}{\mu}\right) IV_k\right] \qquad (10)$$

Therefore, in the marginal probabilities, the parameters associated with upper-level utility is identified upto a factor of λ.

Since the conditional probabilities take the MNL form, we have

$$\pi(j|m) = \left[\exp\left(\mu U_{mj}\right)\right] / \sum_{\forall i} \left[\exp \mu U_{mi}\right] \qquad (11)$$

In equation (11), parameters associated with the lower level alternative is identified up to a factor of μ. The ratio $\frac{\lambda}{\mu}$ is known as the inclusive value parameter. Since the scaling parameter is inversely related to the variance, $\frac{\lambda}{\mu} > 0$, and the previous restriction that $\lambda < \mu$ implies that $0 < \frac{\lambda}{\mu} < 1$.[9]

[9]The condition that the IV parameter must lie within the interval $(0, 1)$ for *a priori*

In the above discussion of nesting structure, only the ratio $\frac{\lambda}{\mu}$ $\left(\text{or } \frac{\lambda^m}{\mu^m}\right)$ is econometrically identified. In the literature, this is done by setting one of the parameters to unity. When we normalize $\mu = 1$, this is named as Random Utility Model 1 (RU1), and when we normalize $\lambda = 1$, this is named as Random Utility Model 2 (RU2) by Hensher and Greene [27]. A minor modification of RU2 is also known as RU3 by them. Most applications of health care demand analysis typically use RU2 version, with the implicit presumption that empirical results are identical to RU1 even though they are numerically different. For a two level nested logit model, invariance across normalizations can be achieved only if $\lambda_1 = \lambda_2$ in RU1 and $\mu_1 = \mu_2$ in RU2 after accounting for scaling.

The non-normalized nested logit (NNNL) on the other hand slightly differs and additional restrictions are required in order for this to be consistent with the utility maximization hypothesis. The marginal choice probabilities in the NNNL model is

$$\pi(m) = \exp\left[U_m + \theta^m IV_m\right] \Big/ \sum_{\forall k} \exp\left[U_k + \theta^k IV_k\right]. \qquad (12)$$

specification of the utility maximizing hypothesis is known as the famous Daly-Zachary-McFadden Condition in the literature. However, this condition has been relaxed for local consistency (Borsch-Supan [8]). This restriction may not ensure RUM to be globally concave, but locally concave over certain ranges.

The conditional choice probabilities in the NNNL model is

$$\pi(j|m) = [\exp(U_{mj})] \Big/ \sum_{\forall i} [\exp U_{mi}] \,. \tag{13}$$

It is shown that the NNNL structure produces identical results only when $\theta^m = \theta^k$, that is the IV parameters must be restricted to equality, and the results will be identical to RU1.

Up until now, we assumed that there are only non-degenerate branches in the nesting structure. If there are one or more degenerate branches (branches with only one alternative, that is Figure2(b) and Figure2(c)), we have additional problems with regard to identification, scaling, and invariance. Hunt [28] argues that a model of type Figure 2(b) or Figure 2(c) in the appendix is overparameterized unless additional restrictions are imposed. In case of RU2 or RU3, the IV parameter associated with a degenerate branch is not an econometrically identified parameter because the IV parameter for the degenerate branch cancels with the scaling parameter. So, the degenerate partition IV parameter must be set to unity trivially. The RU1 allows free parameters in both the branches, but the scaling in the self branch is not identified or estimable econometrically. We estimated both the restricted and unrestricted cases. The NNNL model for the partial degenerate case has an IV parameter that is identified and estimable.

5 Data Source and Variable Construction

The data source comes from the 52nd round of the National Sample Survey (NSS) data, a nationally representative survey conducted by the National Sample Survey Organization (NSSO), Ministry of Statistics, Government of India during July 1995 - June 1996. The survey covered the curative aspects of the general health care system in India including morbidity and utilization of medical services, expenditure incurred for treatment of ailments, utilization of maternity and child health care services, and problems of the aged persons. In this paper, we are analyzing household responses relating to the non-hospitalized ailments and treatment decisions. The NSS survey also contains a wealth of socioeconomic and demographic information reflecting upon individual, household, and community level characteristics.

The NSSO adopted a two stage stratified sampling design. The first stage units were census villages in rural areas and NSSO urban blocks in urban areas; the second stage units were households in both rural and urban areas. The sample villages and urban blocks were selected with a probability proportional to population size in the form of two independent interpenetrating sub-samples. For the selection of households, the frame consisted of three second-stage strata in which a total sample of 10 households was selected. The composition of second stage stratum consist of 2 households reporting at least one child of age '0' year, 2 households reporting any case of hospitalization, and 6 remaining households.

The number of households surveyed on health care in rural and urban areas were 71,284 and 49,658, spread over a sample of 7,663 villages and 4,991 urban blocks, respectively. The survey covered all of India except some interior areas of Nagaland, Andaman & Nicobar Islands, and the Ladakh, Kargil and Dodha districts of Jammu & Kashmir. The number of persons reporting non-hospitalized ailments during 15 days preceding the survey date in rural and urban areas were 21,732 and 13,675, respectively. Regarding non-hospitalized ailments, the survey collected information about the details of health care received, provider-choice, and the amount of out-of-pocket spending for the chosen provider choice.

The outpatients in this paper include all those cases of non-hospitalized ailments reported during 15 days preceding the date of survey. The NSS data is supplied in the form of responses pertaining to different segments of the questionnaire separately without an identification number. In order to link different segments of individual and household responses, we created a household identifier (using the information on state/region, village/block serial number, and second stage stratum number) and an individual identifier (using the household identity and the serial number of the members of a household).

Health Facilities: Health facility alternatives are grouped into two major categories, namely, self care and formal care. Formal care is classified into public facility, private facility, and private doctor. Public facility includes government hospitals, clinics, dispensaries, PHCs, CHCs, and the Central & State government aided ESI facilities. Since various types of public facilities are not numerically important, we merged them into one group, known as government facility. Similarly, private hospitals, nursing homes, and charitable institutions are merged into private hospital group.

Income: The income variable used in the estimation is the monthly household expenditures. Household expenditures are often used in the literature as a good proxy for permanent income. We used this measure of household income because no one in the family is usually denied health care because of her/his income, especially children. We would like to emphasize that in health care choice, the relevant concept is the family.

Price Variables: In order to estimate the model, data on prices (including all medical and non-medical) for all alternatives must be available. As in the case of all previous studies, getting price data posed a difficult problem. The NSS data collected price information (expenditure incurred) for the provider from which the individual received care. In other words, price data were available only for the alternative they chose. Price data for all unchosen alternatives in the choice set for an individual must be obtained in order to estimate a discrete choice model.[10] There is also a question about whether a particular alternative is part of an individual's choice set or not. In order to get around this problem, we adopted a choice generating process and used random sampling with replacement techniques to impute the price of relevant unchosen alternatives in the individual choice set.

After deleting all non-response and not applicable observations, the entire sample were split into small data sets on the basis of province, illness groups (illness on the basis of more than one percent of the frequency; and

[10]In fact, identification of a discrete choice model requires variation across alternatives. Although variation across individulas (such as socio demographic variables) is not necessary, it is desirable to include them in order to obtain precise estimates.

the rest are grouped into other acute and chronic categories, see Table 6 for details), and 5 income quintiles. Each small data set is basically a representative of a geographical region, a specific-illness type, and an income quintile. Within each group, individuals are assumed to be similar but they choose different alternatives; if an alternative is not chosen by anyone in the group then we consider that alternative irrelevant for persons belonging to that particular group. This process generated the first stage of a differential choice set. In the second phase, for an individual's unchosen alternative in the choice set, we draw 10 random draws of price (medical expenditures) with replacement from those within the group who choose that alternative. The mean of this random sample is the imputed price for the unchosen alternative. This process is repeated for every individual within all groups to generate the relevant price data. In the second stage, if the imputed price exceeds family income, that alternative is considered to be not feasible and hence dropped from the individual's choice set.[11] In this way, we embedded heterogeneity arising from geographical regions, illness types, income groups, and individual idiosyncratic elements into our model that resulted a variable choice set scenario coupled with a heterogeneous price variable.

[11] The STATA syntax for imputation and generation of a variable choice set are availbale upon request.

Quality: Like many other previous studies, obtaining data on quality of provider choice is problematic. We use the number of treatments received as an indicator of quality. Like the price variable, the number of treatments is available conditional on choosing an alternative. We resorted to the same random sampling imputation procedures to arrive at the expected number of treatments for all unchosen alternatives in the choice set.

Distance: Another important variable which varies across alternatives that is available in rural sample is the distance. The distance variable is not available in the survey data but we extracted this information from the sample village characteristics provided by the NSSO.

All other variables used in the estimation are individual-specific or household specific variables, such as age, education, household size, drinking water facility, latrine facility, etc. Table 5 provides all variable definitions and Table 7 shows the descriptive statistics of variables for the rural sample in the estimation.

6 Estimated Results

The parameters of NMNL models are estimated by using the full information maximum likelihood method of NLOGIT. We also estimated a MNL model for comparison. In a discrete choice model, a reference alternative must be identified because only a difference in utility between an alternative and a reference alternative matters in the estimation. As in previous studies, the coefficients of the log of net consumption and its square are assumed to be constant across choices and we estimated both the nesting structures. We also constrained all distance variables to be the same for all choices but the coefficient on treatments is allowed to vary across alternatives. At this point, we have not added individual specific and household specific variables for the sake of model comparisons. The estimated results are reported in Table 8.

It can be seen from Table 8 that RU1 and NNNL models are numerically identical when IV parameters are restricted to equality. However, with IV parameters unrestricted, the estimates are not invariant across scaling normalizations. This means that there is no obvious relationship between these two parameter estimates, and the log-likelihood values at convergence are not identical (-19269.62 vs. -19419.66 in NMNL1 or -19268.33 vs. -19425.48 in NMNL2). This is a problem because there can be only one utility maximizing model. Invariance can be attained by adding appropriate dummy nodes[12] or restricting the IV parameters to equality.

In this paper, for NMNL models with IV parameters restricted to equality, the reported result is an estimate of λ when $\mu = 1$ is the scaling normalization; and an estimate of μ (not $\frac{1}{\mu}$) when $\lambda = 1$ is the scaling normalization.

When we estimated RU1 (or NNNL) and RU2 with restrictions that the IV parameters be equal in the former and the degenerate partition IV parameter set to unity in the later, invariance is achieved across normalizations after accounting for scaling. The log-likelihood functions are equal (-19419.66 in NMNL1 or -19425.48 in NMNL2) and the IV parameters are inverse to one another $\left(\frac{1}{0.82} = 1.221\,5 \text{ or } \frac{1}{0.89} = 1.125\,3\right)$. Multiplying the utility function parameter estimates by the corresponding IV parameter estimates produce equivalent results within rounding errors.

The coefficients on log consumption and log consumption squared are statistically significant, which implies that price and income do play an important role in the demand for outpatient health care. At this point, we added individual specific and household specific variables into our model so as to obtain precise estimation. The estimated results of this complete spec-

[12]See Koppelman and Wen [31], and Hensher and Greene [27] for details about adding dummy nodes.

ification is reported in Table 9 for MNL and RU2, respectively.

Since price and income enter the model in a highly non-linear fashion, it is difficult to make any assessment by just looking at these coefficients. So, we simulated predicted probabilities at different price levels in order to make an assessment of the impact of price on demand for outpatient care in rural India and summarized the results across income quintiles. The simulated results for change in price from Rs. 25 to Rs. 200 and the the corresponding predicted demands for government facility, private facility and doctor's clinic across income quintiles are reported in Tables 10(a) - 10(c), respectively. The corresponding graphs depicted in Figures 3(a) - 3(c). In both MNL and NMNL specifications, the own price effects are negative everywhere indicating that demand curves are downward sloping and cross price effects are mostly positive. If price is too high, poor will demand relatively less health care from the formal sector than the rich. Further, an increase in private doctor's fees shifts utilization towards self care more than increase in price of of care available at government facilities.

However, the overall effect of price on demand is relatively inelastic, which indicates that there is a potential increase in revenue if user fees are imposed for publicly provided outpatient care. But then, as discussed earlier, the demand for health care is related to the health seeking behaviour of the people which in turn in is embedded in social and cultural aspects of viewing morbidity, reporting illness behaviour, and choice of treatment. If that is so, imposing user fees of both direct and indirect nature may not be welfare enhancing and poor people will bear the burden of health. This is because more spending on health care means less resources are available for other consumption goods and services. The outcome may be worse for the poor since their reporting illness behaviours are such that they seek care for catastrophic illnesses only. Since the price elasticity of demand is low, this suggests that private providers can effectively raise the price of care without losing their potential customers.

The coefficients on distance variables are all negatively significant, implying that the higher the distance to a formal health care facility the lower the demand for health care. Moreover, higher the distance, the effect on reducing demand is higher in magnitude. This suggests that non-monetary factors do play an important role in the demand for health care.

Quality of provider choice, proxied by the number of treatments, is positive and statistically significant determinant of choosing a private hospital, whereas this coefficient is negatively significant for private doctors and statistically insignificant for government facility. As an alternative interpretation, the higher number of treatments in private hospitals could be a possible evidence of induced demand.

Age has a positive effect on the use of government facility type and private hospital. The effect of age on doctor's clinic is negative. The effect of age squared is negative for all types of formal health care alternatives. This suggest that older patients prefer government facilities and private hospitals. Moreover, the illnesses suffered by the older patients can be treated in hospitals rather than doctor's clinc.

Females are less likely to use formal care than males, suggesting possible gender bias in health care utilization. There could be two explanations for this findings. In terms of human capital theory argument, households may invest more on more productive members of the household. Since most females in rural India engage in household chores, females may be viewed as less productive members of the household. The second argument is cultural biasness towards females in rural India. For instance, many poor women, especially those from lower castes, participate in the labour market but they do not get equal treatment when it comes to household resources. Similarly widows do not inherit family assets and wealth in rural India. However, possible gender bias towards formal health care utilization in rural India could mainly be attributed to cultural biasness towards females.

Educated individuals use more formal care of all types. So, the conventional argument holds that education increases the expected productivity of formal health care alternatives relative to self care. This is also in consonance with the general notion that the pattern of reporting morbidity and contacting a health professional tends to increase with the level of education. The effect of household size on the demand for health care is positive and significant. More members of the household may imply less attention to members of the household in terms of their appropriate nutritional intakes thereby contacting illnesses and utilization of more formal care. Alternatively, the reporting behaviour may be improving with large families due to higher family disposable income.

Lack of access to safe drinking water at the household level has negative effects on the demand for health care. The number of bad habits has also negative effect on the demand for health care. Also, those households who do not have latrine facilities are less likely to contact private facilities or doctor's clinic. These results seem to be somewhat paradoxical at the first instance. But a close look at the Indian society at the grass root level does reflect this pattern of behaviour. People in extreme poor and poverty have a lack of access to basic needs, such as adequate food, clothing, and a pucca house. So, obtaining health care from the formal sector in the event of minor illnesses is really a luxury for them. My personal experience shows that poorer people tend to obtain care from the formal sector only in the event of catastrophic illnesses, subject to availability of money from local money lenders or disposing off household assets. This does not mean that poor people do not seek care for minor illnesses; they do seek care from the informal sector, such as traditional healers, quackes, etc. that are relatively cheaper and available around local areas.

7 Conclusions

In this paper, we estimated nested multinomial logit models of the demand for outpatient health care in rural India using the most recent NSSO data. We extend the conventional empirical modelling approach in the context of a typical developing country like India. This paper incorporated a range of heterogeneous elements ranging from geographical location to individual idiosyncractic elements inevitable in the household data. Various elements of heterogeneity are reflected through a variable choice set, the number of expected treatments received, and the expected price of care at the point of access. Contrary to some of the previous studies, we found that prices and income embedded in the log of net expenditure variables are statistically significant determinants of health care choice in rural India. Although the estimated price elasticities are small, they are higher for lower income groups than higher income groups. Distance is one of the most pronounced inhibiting factor in the demand for health care. Another set of inhibiting factors governing health seeking behaviour is those which are attributed to stylized facts of the socioeconomic environment - especially those who have bad habits and those who do not have access to safe drinking water and latrine facilities. The result is suggestive of gender bias in the demand for health care.

One of the major weakness of this study is that the analysis was restricted by the available data quality. So, our results should be interpreted with great care. In order to effectively use a discrete choice model of health care demand, it is important that data on quality and other characteristics of alternatives need to be collected along with the survey data. Although we tried with different imputed data sets and obtained similar results, the estimated results would greatly improve with precision in the presence of a high quality data. Nevertheless, the present study is a preliminary attempt in the direction of modelling demand for health care in rural India and can motivate future research.

At the technical level, there could be a sample selection bias problem. The information on health care decision is reported conditional on reporting an illness within 15 days. However, for given health status, rich and poor may make different reporting patterns, and other knowledge and information that attribute to morbidity may be correlated with some of the regressors in the model. It is because of these peculiar features, using a sample of those who report an illness may result in selectivity bias. In particular, sample selection bias arises if unobserved individual characteristics simultaneously determine morbidity and choice of treatment.

Although this issue has been recognized in many studies, most papers ignored the problem and treat the sample of those who report illness as a subsample. A two-step procedure for the discrete choice model similar to Heckman was suggested by Van de Ven and Van Praag [47] in the context of a probit model. However, this approach has been difficult to implement because of absence of proper instruments that affect illness but not the choice of treatment. In fact, it is perhaps because of poor instruments in the data, those studies which did attempt to correct for sample selection bias did not find any statistical selection bias. In the light of inconclusive evidence about the selection bias, absence of suitable instruments, and the computational complexities of multinomial probit model, the nested multinomial logit models without controlling for sample selection bias may be appropriate.

References

[1] Acton, P. (1975),"Non-monetary factors in the demand for health services: some empirical evidence" *Journal of Political Economy*, 83(3), 595-614.

[2] Akin, J. S., Griffin, C., Guilkey, D. K., and Popkin, B. M. (1984), "The Demand for Primary Health Care in The Third World", Totowa, N. J., Littlefield, Adams.

[3] Akin, J. S., Griffin, C., Guilkey, D. K., and Popkin, B. M. (1986), "The demand for primary health care services in the Bicol region of the Phillippines", *Economic Development and Cultural Change* **34(4)**, 755-782.

[4] Ben-Akiva, M., and Lerman, S. R. (1985), "Discrete Choice Analysis: Theory and Application to Travel Demand", Cambridge, MA, The MIT Press.

[5] Birdshall, N. and Chuhan, P. (1986), "Client choice of health treatment in rural Mali", *Population and Human Resources Department*, World Bank, Washington, D. C.

[6] Blomqvist, A., Wang, S., and Zhu L. (1998), "Financing health care in developing economies: a simulation approach", University of Western Ontario, London, Ontario.

[7] Bolduc, D., Lacroix, G., and Muller, C. (1996), "The choice of medical providers in rural Bénin: a comparison of discrete choice models", *Journal of Health Economics*, 15, 477-498.

[8] Borsch-Supan, A. (1990), "On the compatibility of nested logit models with utility maximization", *Journal of Econometrics,* 43, 373-388.

[9] Cameron, A. C. Trivedi, P. K., Milne, F., and Piggot, J. (1988), "A microeconometric model of the demand for health care and health insurance in Australia", *Review of Economic Studies,* 55(1), 85-106.

[10] Dor, A., Gertler, P., and Van der Gaag, G. (1987), "Non-price rationing and the choice of medical care providers in rural Cote D'Ivoire", *Journal of Health Economics*, **6**, 291-304.

[11] Dor, A. and van der Gaag, G. (1993), "Quantity rationing and the demand of adults for medcal care in rural Cote d'Ivoire ", in Health Economics Research in Developing Countries A. Mills and K. Lee (eds.), Oxford University Press.

[12] Dow, W. H. (1995), "Discrete choice estimation of price-elasticities: the benefits of a flexible behavioral model of health care demand", Discussion paper No. 739, Economic Growth Center, Yale University, New Haven, CT.

[13] Dow, W. H. (1999), "Flexible discrete choice demand models consistent with utility maximization: an application to health care demand", *American Journal of Agricultural Economics*, 81, 680-685.

[14] Duggal, R, and Amin, S. (1989), "Cost of health care: a household survey in an Indian district", The Foundation Research in Community Health (FRCH), Bombay, India.

[15] Ellis, R. P., McInnes, D. K. and Stephenson, E. H. (1994), "Inpatient and outpatient health care demand in Cairo, Egypt", *Health Economics,* 3, 183-200.

[16] Greene, W. H. (1997), "Econometric Analysis", 3rd edition, *Prentice Hall.*

[17] Grrenhalgh, T. (1987), "Drug prescription and self-medication in India: an exploratory survey", *Social Science and Medicine*, 25, 307-18.

[18] Gertler, P., Locay, L., and Sanderson, W. (1987), "Are use fees regressive? The welfare implications of health care financing proposals in Peru" *Journal of Econometrics*, 36, 67-88.

[19] Gertler, P., and van der Gaag, J. (1990), "The willingness to Pay for Medical Care: Evidence from two Developing Countries", Baltimore, The John Hopkins University Press.

[20] Gertler, P. and Hammer, J. (1997), "Strategies for pricing publicly provided health services", World Bank Policy Research Working Paper 1762.

[21] Grossman, M. (1972a), "On the concept of health capital and the demand for health", *Journal of Political Economy*, 80(2), 223-235.

[22] Grossman, M. (1972b), "The demand for health - a theoretical and empirical investigation", (NBER) New York, Columbia University Press.

[23] Heller, P. (1982), "A model of the demand for medical and health services in Peninsular Malaysia", *Social Science and Medicine*, 16, 267-84.

[24] Hajivassiliou, V. A., MacFadden, D., and Ruud, P. (1996), "Simulation of multivariate normal rectangle probabilities and their derivatives - theoretical and computational results", Journal of Econometrics, 72, 85-134.

[25] Hausman, J., and McFadden, D. L. (1984), "Specification tests for the multinomial logit model", *Econometrica*, 52(5), 1219-1240.

[26] Hausman, J., Leonard, G. K., and McFadden, D. L. (1995), "A utility consistent, combined discrete choice and count data model assessing recreational use losses due to natural resource damage", *Journal of Public Economics*, 56, 1-30.

[27] Hensher, D. A., and Greene, W. H. (2002), "Specification and estimation of nested logit model: alternative normalizations", *Transportation Research*, 36(1), 1-17.

[28] Hunt, G. L. (2000), "Alternative nested logit model structures and the special case of partial degeneracy", *Journal of Regional Science*, 40(1), 89-113.

[29] Jack, W. (1999), "Principles of Health Economics for Developing Countries", WBI Development Studies, World Bank.

[30] Jiminez, E. (1995), "Human and physical infrastructure: public investment and pricing policies in developing countries" in T. N. Srinivasan and J. Behrman (eds.), Handbook of Development Economics, Volume 3, Amsterdam, North Holland.

[31] Koppelman F. S., and Wen, C. H. (1998), "Alternative nested logit models: structure, properties and estimation", *Transportation Research* 32B(5), 289-298.

[32] Krishnan, T. N. (1995), "Access and burden of treatment: an inter-state comparison", Centre for Development Studies, Thiruvanathapuram, India.

[33] Manski, C. F. and McFadden, D. (1981), "Alternative estimators and sample designs for discrete choice analysis", in *Structural Analysis of Discrete Data.* Edited by C. F. Manski and D. McFadden. Cambridge, MIT Press.

[34] McFadden, D. L. (1974), "Conditional logit analysis of qualitative choice behavior", in *Frontiers in Econometrics* edited by Paul Zarembka, New York, Academic press.

[35] McFadden, D. L. (1981), "Econometric models of probabilistic choice", in *Structural Analysis of Discrete Data.* Edited by C. F. Manski and D. McFadden, Cambridge, MIT Press.

[36] McFadden, D. L. (1989), "A method of simulated moments for estimation of discrete response models without numerical integration", *Econometrica,* 57(5), 995-1026.

[37] Mwabu, G. (1986), "Health care decisions at the household level: results of health survey in Kenya", *Social Science and Medicine*, 22(3), 313-319.

[38] Mwabu, G., Ainsworth, M., and Nyamete, A. (1993), "Quality of medical care and choice of medical treatment in Kenya: an empirical analysis", *The Journal of Human Resources*, XXVIII (4), 838-862.

[39] Mwabu, G., Mwanzia, J., and Laimbila, W. (1995), "User charges in government health facilities in Kenya", *Health Policy and Planning*, 10(2), 164-170.

[40] NCAER (1993), "Household survey of health care utilisation and expenditure", *Working Paper #53*, New Delhi, India.

[41] NSSO (1998), "Morbidity and treatment of ailments", *NSSO Report Number 441*, Department of Statistics, Government of India.

[42] Phelps, C. E. (1992), "Health Economics", New York, Harper Collins.

[43] Sarma, S. K. (1995), "Access to health care: a socioeconomic study of two locations in the Ganjam District of Orissa", *unpublished dissertation submitted to the Jawaharlal Nehru University*, New Delhi, India.

[44] Schwartz, J. B., Akin, J. S., and Popkin, B. M. (1988), "Price and income elasticities of demand for modern health care: the case of infant delivery in the Phillippines", *World Bank Economic Review*, **2(1)**, 49-76.

[45] Uplekar, M. (1989a), "Implications of prescribing patterns in private doctors in the treatment of pulmonary tuberculosis in Bombay", *Report*

#40, Takemi Program in International Health, Harvard School of Public Health.

[46] Uplekar (1989b), "Private doctors and public health: the case of leprosy in Bombay",*Report #41, Takemi Program in International Health*, Harvard School of Public Health.

[47] Van de Ven, W. P. M. M. and Van Praag, B. M.S. (1981), "The demand for deductibles in private health insurance", *Journal of Econometrics*, 17, 229-252.

[48] Vishwanathan, H. and Rhode, J. (1990), "Diarrhoea in rural India: a nation-wide study of mothers and practitioners ", Vision Books, New Delhi, India.

[49] Yesudian, C. (1994), "The nature of private sector health services in Bombay", Health Policy and Planning, 9(1), 72.

[50] Zweifel, P. and Bryer, F. (1997), "Health Economics", Oxford University Press, New York.

Figure1: Relationship between health and health care

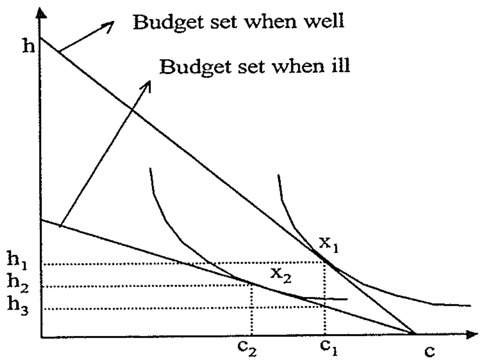


Figure 1a

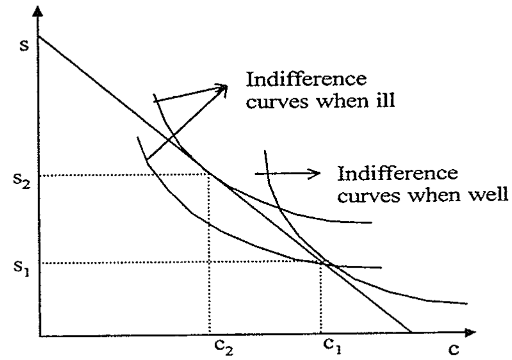


Figure 1b

Source: William Jack (1999)

Figure 2: Decision Trees

Provider Choice Model (MNL)

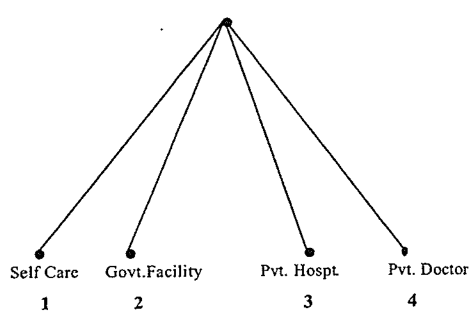


Figure 2 (a): Complete Degeneracy

Provider Choice Model (NMNL 1)

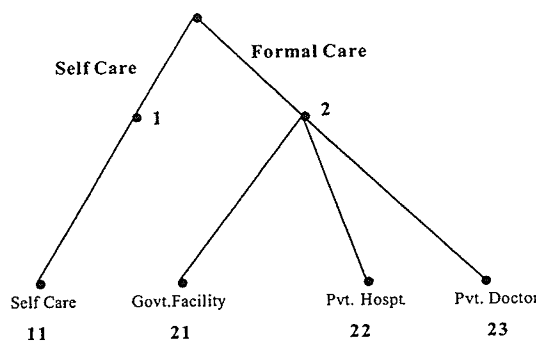


Figure 2 (b): Two Level Nested Structure with one Partial Degenerate Branch

Provider Choice Model (NMNL 2)

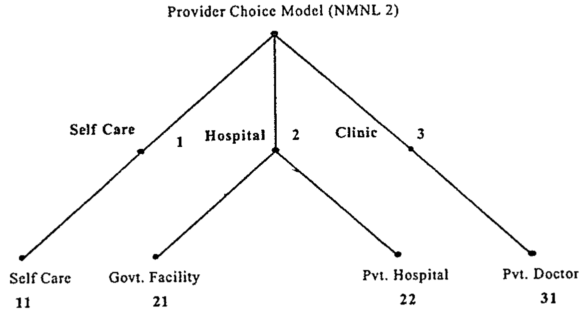


Figure 2(c): Two Level Nested Structure with two Partial Degenerate Branches

Table 1
Goals and Achievements for Health and Family Welfare Programmes in India

	Indicator	Initial Level	GOALS 1985	1995	2000	Achievements
1	Infant Mortality Rate Prenatal morality	Rural: 136 (1978) Urban: 70 (1978) Total: 125 (1978) 67 (1976)	122 60 106	 87	 < 60 30-35	 72 (1998) 64*** (1990)
2	Crude death rate	Around 14	12	10.4	9.0	9.0 (1998)
3	Pre-school Child mortality (1-5 years) **CMR (0 to 4 years)**	24 (1976-77) **42.2 (1984)**	20-24	15-20	10	N.A. **24.2**
4	Maternal mortality rate	4-5 (1976)	3-4	2-3	<2	570 (ratio)
5	Life expectancy at birth (years)	Male: 52.6 (1976-81) Female: 55.6 (1976-81)	55.1 54.3	57.6 57.1	64 64	62.1** 62.7**
6	Babies with birth weight below 2500 grams (in percentage)	30	25	18	10	30 (1992)
7	Crude birth rate	Around 35	31	27.0	21.0	26.4 (1998)
8	Effective couple protection (percentage)	23.6 (1982)	37.0	42.0	60.0	46.5*
9	Net reproduction rate	1.48 (1981)	1.34	1.17	1.00	N.A.
10	Growth rate (annual)	2.24 (1971-81)	1.90	1.66	1.20	1.74 (1998)
11	Family size	4.4 (1975)	3.8		2.3	
12	Pregnant mothers receiving ante-natal (%)	40-50	50-60	60-75	100	
13	Deliveries by trained attendants (%)	30-35	50	80	100	44.1
14	Immunization status (% coverage)	TT for pregnant women 20 TT for school children: 10 years 16 years 20 DPT (children < 3 yrs) 25 Polio (infants) 5 BCG (infants) 65 DT (new school entrants) 20 Typhoid (......) 2	60 40 60 70 50 70 80 70	100 100 100 85 70 80 85 85	100 100 100 85 85 85 85 85	75.06 (96-97) 89.73 (96-97) 87.21 (96-97) 93.68 (96-97)
15	Leprosy: percentage of arrested cases out of those detected	20	40	60	80	estimated cases (1995) 951,500
16	TB: percentage of disease arrested cases out of those detected	50	60	75	90	No. of cases 226,543 (1994)
17	Blindness incidence (percentage)	1.4	1	0.7	0.3	

* Provisional figures from Sample Registration System, 1996 & 1998; ** The State of World Population 1997 (UNFPA Publication); *** World Bank (1993).

Table 2
Incidence of Ailments During the last 15 days

Area	Ailment	Male		Female		Total	
		Sampled	Estimated (00)	Sampled	Estimated (00)	Sampled	Estimated (00)
Rural	Acute	8,191	1,34,117 **(41)**	8,320	1,35,656 **(44)**	16,511	2,69,773 **(42)**
	Chronic	2,692	41,773 **(13)**	2,629	42,528 **(14)**	5,321	84,301 **(13)**
	Any	10,832	17,5224 **(54)**	10,900	17,7401 **(57)**	21,732	3,52,625 **(55)**
Urban	Acute	4,934	41,120 **(39)**	4,921	41,592 **(43)**	9,855	82,712 **(41)**
	Chronic	1,857	13,297 **(13)**	2,005	14,844 **(15)**	3,862	28,141 **(14)**
	Any	6,767	54,264 **(51)**	6,908	56,263 **(58)**	13,675	1,10,527 **(54)**

Figures in parenthesis are per 1000 persons. Acute ailments refer to short duration (less than 30 days) ailments and chronic ailments refer to long duration ailments (30 days or more).
Source: NSSO (1998).

Table 3
Distribution of Treatments During last 15 days by Source of Treatment

Source	Rural		Urban	
	Sampled	Estimated (00)	Sampled	Estimated (00)
Public Hospital	2,184	29,898 **(101)**	1,902	14,070 **(136)**
PHC/CHC	1,065	15,380 **(52)**	122	706 **(7)**
Public Dispensary	387	5,748 **(19)**	163	1,486 **(14)**
Private Hospital	1,957	32,745 **(110)**	1,945	1,806 **(143)**
Nursing Home	448	7,237 **(24)**	363	2,279 **(22)**
Charitable Institution	77	1,013 **(3)**	102	785 **(8)**
ESI Doctor/AMA, etc.	44	782 **(3)**	106	945 **(9)**
Private Doctor	8,112	1,44,903 **(488)**	5,966	51,848 **(502)**
Others	1,656	25,219 **(85)**	863	6,841 **(66)**
Not Reported	2,181	34,268 **(115)**	1,157	9,426 **(91)**
Total	18,110	2,97,193	12,662	1,03,192

Source: NSSO (1998). Figures in parenthesis are per 1000 persons.

Table 4
Average Medical and Non-medical Expenditure per Treated Ailment During 15 Days by Source of Treatment

Type	Source	Rural	Urban
Medical Expenditure (In Rupees)	Government	110	146
	Other	168	185
	All	157	178
Non-Medical Expenditure (In Rupees)	Government	19	20
	Other	18	15
	All	19	16
Total Expenditure (In Rupees)	Government	129	166
	Other	186	200
	All	176	194

Source: NSSO (1998).

Table 5
Variable Definitions

Variable	Definition
Government Facility	Includes the following: Public Hospital, Primary Health Care Centre, Public Dispensary, and ESI Doctor.
Private Facility	Includes the following: Private Hospital, Nursing Home, Charitable Institutions, and Other Medical Institutions.
Private Doctor	Registered individual practitioners.
Self Care	Includes the following: self-medication, or advice from other household members, friends, medicine shop, and other non-medical professional practitioners.
Price (in Rupees)	Includes the following items within the reference period: a) Total medical expenditure incurred for treatment; b) Transport and Lodging; c) Personal medical appliances; d) Any reimbursement by employer or other agencie s; e) Loss of household income; and f) Other expenditure.
Income	Monthly household consumption expenditure (includes average monthly medical expenses and consumer durables).
Distance	= 1 if the distance to the nearest facility is located within less than two kilometers from the sample village, otherwise = 0.
Distance a	= 1 if the distance to the nearest facility is located within two kilometers to less than five kilometers from the sample village, otherwise = 0.
Distance b	= 1 if the distance to the nearest facility is located within five kilometers to less than ten kilometers from the sample village, otherwise = 0.
Distance c	= 1 if the distance to the nearest facility is located within ten kilometers or more from the sample village, otherwise = 0.
Treatments	Number of treatments. This includes the following: a) Drugs or preparations used for treating an ailment; b) X-ray, ECG (electro-cardiogram), and ECG (electro-encephalogram); c) Other diagnostic tests; d) Surgical operations; and e) Any other treatment.
Age	Age in years.
AgeD	$(\text{Age} - \text{Mean}_{age})^2$
Female	Female=1, Male=0.
Education	= 1 if middle level or above, otherwise = 0.
Household Size	Number of household members.
SCST	= 1 if Scheduled Caste or Scheduled Tribe (socially and economically backward), otherwise = 0.
Habits	Number of habits. This includes regular habit of consumption of the following items: a) Alcohol; b) Biri/Cigar/Cigarette/Hukka; c) Tobaco; d) Ganja; e) Charas; and f) Opium.
Drinking Water	= 1 if the source of drinking water for the household is unsafe: river, canal, other sources; 0 if the source of drinking water for the household is safe: tap, tube-well/hand pump, tankers, pucca well, tank/pond reserved for drinking water.
Latrine	= 1 if there is no latrine facility availed by the household or the type of latrine is non-septic; 0 if the type of latrine facility availed by the household is service latrine or septic tank or flush system.

Table 6
Distribution of Ailments in Rural Area (Outpatient Care)

Name of Ailment	Code	Frequency	Percent
Fevers of short duration	106	6755	37.11
Other diagnosed acute ailment	198	2066	11.35
Diarrhoea & gastro-entertitis dysentery	101	1317	7.23
Cough and acute bronchitis	114	1226	6.73
Other diagnosed chronic diseases	298	996	5.47
Pain in the joints	233	537	2.95
Other undiagnosed acute ailment	199	447	2.46
Injury due to accident and violence	118	404	2.22
Gastritis and hyper-acidity	229	327	1.8
High or low blood pressure	225	322	1.77
Pulmonary tuberculosis	202	300	1.65
Acute respiratory infection	115	265	1.46
Whooping Cough	104	257	1.41
Diseases of heart	224	208	1.14
Diseases of mouth, teeth, and gum	116	203	1.12
Diseases of eye	110	201	1.1
Other undiagnosed chronic diseases	299	154	0.85
Diseases of kidney or urinary system	230	152	0.83
Diabetes	212	139	0.76
Locomotor disability	235	133	0.73
Other disorders of bones and joints	234	132	0.73
Mental and other behavioural disorders	216	115	0.63
Visual disabilities (excludes cataract)	219	112	0.62
Chicken pox	107	111	0.61
Pregnancy complications	117	97	0.53
Other disease of Nerves	218	92	0.51
Acute disease of ear	111	92	0.51
Cataract	220	87	0.48
Other diseases of the eye	221	79	0.43
Hearing disability	222	63	0.35
Cancer	208	64	0.35
Measles or German Measles	108	63	0.35
Other tumours	209	54	0.3
Jaundice	205	53	0.29
Piles	226	49	0.27
Chronic ameobiosis	201	48	0.26
Meningitis & viral encephalitis	105	45	0.25
General debility anaemia	210	44	0.24
Epilepsy	217	41	0.23
Other diseases of the ear	223	39	0.21
Mumps	109	38	0.21
Leprosy	203	30	0.16
Heart failure	112	29	0.16
Filaria (elephantisis)	207	27	0.15
Diseases of mouth, teeth, and gum	228	25	0.14
Cerebral stroke	113	26	0.14
Goitre & thyroid conditions	211	24	0.13
Speech disability	227	22	0.12
Diptheria	103	16	0.09
Other mulnutrition diseases	215	15	0.08
Ricket	214	14	0.08
Tetanus	102	12	0.07
Hydrocele	232	10	0.05
Other congenital deformities	236	8	0.04
Sexual transmitted diseases	204	7	0.04
Prostate disorders	231	6	0.03
Beri beri	213	4	0.02
Guinea worm	206	3	0.02
Total		18205	100.02

Table 7
Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Government Facility[a]	14648	0.1878	0.3905	0	1
Private Facility[a]	15026	0.2180	0.4129	0	1
Private Doctor[a]	15837	0.4587	0.4983	0	1
Self Care[a]	14148	0.238	0.4262	0	1
Price (in Rupees)[1]	14648	205.9702	293.4382	1	6534.5
Price (in Rupees)[2]	15026	250.3908	342.5602	1	6407.5
Price (in Rupees)[3]	15837	206.3832	261.4856	1	3800
Price (in Rupees)[4]	14148	0	0	0	0
Distance[1]	14648	0.0661	0.2485	0	1
Distance[2]	15026	0.1151	0.3191	0	1
Distance[3]	15837	0.4481	0.4973	0	1
Distance[4]	14148	0	0	0	0
Distance a[1]	14648	0.1222	0.3275	0	1
Distance a[2]	15026	0.1471	0.3542	0	1
Distance a[3]	15837	0.2455	0.4304	0	1
Distance a[4]	14148	0	0	0	0
Distance b[1]	14648	0.2056	0.4041	0	1
Distance b[2]	15026	0.2266	0.4186	0	1
Distance b[3]	15837	0.1721	0.3775	0	1
Distance b[4]	14148	0	0	0	0
Distance c[1]	14648	0.6060	0.4886	0	1
Distance c[2]	15026	0.5111	0.4998	0	1
Distance c[3]	15837	0.1342	0.3408	0	1
Distance c[4]	14148	0	0	0	0
Treatments[1]	14648	1.1897	0.4903	0	5
Treatments[2]	15026	1.14173	0.5633	0	5
Treatments[3]	15837	1.13937	0.4281	0	5
Treatments[4]	14148	0	0	0	0
Income[*]	16668	2199.556	1510.09	151	37266
Age	16668	31.1027	24.2607	0	99
AgeD	16668	589.5555	593.7149	0.01	4747.21
Female	16668	0.5044	0.4999	0	1
Education	16668	0.1352	0.3419	0	1
Household Size	16668	6.2201	3.1334	1	33
SCST	16668	0.3080	0.4617	0	1
Habits	16668	0.3102	0.5878	0	5
Drinking Water	16668	0.0689	0.2535	0	1
Latrine	16668	0.9350	0.2463	0	1

[a] =1, otherwise=0.
[1] Government Facility.
[2] Private Facility.
[3] Private Doctor.
[4] Self Medication or others.
[*] Proxied as monthly household consumption expenditure.

Table 8 (a)
FIML Estimation Results from Alternative Model Specifications

Variable	MNL	NMNL 1 (NNNL)	NMNL 2 (NNNL)	NMNL 1 (IV Restricted) (NNNL)	NMNL 2 (IV Restricted) (NNNL)
Constant[1]	0.46625 (7.091)	1.6149 (2.658)	2.2851 (2.850)	0.2344 (2.931)	0.5788 (6.691)
Constant[2]	0.3794 (6.966)	1.5183 (2.506)	2.1798 (2.726)	0.1517 (2.149)	0.4825 (6.429)
Constant[3]	1.2179 (23.017)	2.3642 (3.90)	2.9162 (3.678)	0.9858 (14.137)	1.3586 (15.447)
Log Consumption[a]	-0.5946 (-2.573)	-0.8505 (-4.361)	-0.9513 (-4.185)	-0.4866 (-2.366)	-0.664 (-2.795)
Log Cons. Squared[a]	0.0782 (4.371)	0.0913 (5.484)	0.1026 (5.258)	0.0685 (4.002)	0.0864 (4.362)
Distance a[a]	-0.2814 (-7.60)	-0.2539 (-7.145)	-0.251 (-5.671)	-0.2684 (-7.643)	-0.3126 (-7.214)
Distance b[a]	-0.517 (-13.579)	-0.4809 (-13.035)	-0.4897 (-9.723)	-0.4948 (-13.577)	-0.5634 (-12.052)
Distance c[a]	-0.8045 (-21.738)	-0.7405 (-20.056)	-0.8073 (-15.313)	-0.7659 (-21.029)	-0.8744 (-17.169)
Treatments[1]	-0.0238 (-0.542)	-0.0393 (-0.915)	-0.0413 (-0.889)	-0.0241 (-0.561)	-0.0275 (-0.599)
Treatments[2]	0.1514 (4.495)	0.1404 (4.265)	0.1554 (4.238)	0.1447 (4.426)	0.159 (4.421)
Treatments[3]	-0.1956 (-4.942)	-0.1975 (-5.152)	-0.2052 (-4.743)	-0.1903 (-5.059)	-0.22 (-4.791)
Inclusive Value					
Self		-0.1018 (-0.528)		1.2215 (22.208)	
Formal		1.1881 (21.014)		1.2215 (22.208)	
Inclusive Value					
Self			-0.2223 (-1.371)		0.8887 (19.256)
Hospital			0.8517 (17.715)		0.8887 (19.256)
Doctor			0.9706 (14.792)		0.8887 (19.256)
Log-likelihood	-19428.0	-19269.62	-19268.33	-19419.66	-19425.48
Observations	16668	16668	16668	16668	16668

[a] The coefficients are restricted to be equal across equations.
[1] Government Facility.
[2] Private Facility.
[3] Private Doctor.
Figures in parentheses are t-statistics.

Table 8 (b)
FIML Estimation Results from Alternative Model Specifications

Variable	MNL	NMNL 1 (IV Restricted) (RU1)	NMNL 2 (IV Restricted (RU1)	NMNL 1 (IV Restricted) (RU2)	NMNL 2 (IV Restricted) (RU2)
Constant[1]	0.46625 (7.091)	0.2344 (2.931)	0.5788 (6.691)	0.2863 (3.460)	0.5144 (8.538)
Constant[2]	0.3794 (6.966)	0.1517 (2.149)	0.4825 (6.429)	0.1853 (2.378)	0.4288 (7.830)
Constant[3]	1.2179 (23.017)	0.9858 (14.137)	1.3586 (15.447)	1.2042 (21.297)	1.2074 (24.312)
Log Consumption[a]	-0.5946 (-2.573)	-0.4866 (-2.366)	-0.664 (-2.795)	-0.5944 (-2.458)	-0.5901 (-2.922)
Log Cons. Squared[a]	0.0782 (4.371)	0.0685 (4.002)	0.0864 (4.362)	0.0836 (4.188)	0.7675 (4.638)
Distance a[a]	-0.2814 (-7.60)	-0.2684 (-7.643)	-0.3126 (-7.214)	-0.3279 (-7.630)	-0.2778 (-7.781)
Distance b[a]	-0.517 (-13.579)	-0.4948 (-13.577)	-0.5634 (-12.052)	-0.6044 (-12.828)	-0.5007 (-13.455)
Distance c[a]	-0.8045 (-21.738)	-0.7659 (-21.029)	-0.8744 (-17.169)	-0.9355 (-18.813)	-0.777 (-21.159)
Treatments[1]	-0.0238 (-0.542)	-0.0241 (-0.561)	-0.0275 (-0.599)	-0.0294 (-0.691)	-0.0244 (-0.733)
Treatments[2]	0.1514 (4.495)	0.1447 (4.426)	0.159 (4.421)	0.1768 (4.876)	0.1413 (4.888)
Treatments[3]	-0.1956 (-4.942)	-0.1903 (-5.059)	-0.22 (-4.791)	-0.2325 (-5.531)	-0.1953 (-5.461)
Inclusive Value					
Self		1.2215 (22.208)		1.00 (fixed)	
Formal		1.2215 (22.208)		0.8186	
Inclusive Value					
Self			0.8887 (19.256)		1.00 (fixed)
Hospital			0.8887 (19.256)		1.1252 (17.927)
Doctor			0.8887 (19.256)		1.00 (fixed)
Log-likelihood	-19428.0	-19419.66	-19425.48	-19419.66	-19425.48
Observations	16668	16668	16668	16668	16668

[a] The coefficients are restricted to be equal across equations.
[1] Government Facility.
[2] Private Facility.
[3] Private Doctor.
Figures in parentheses are t-statistics.

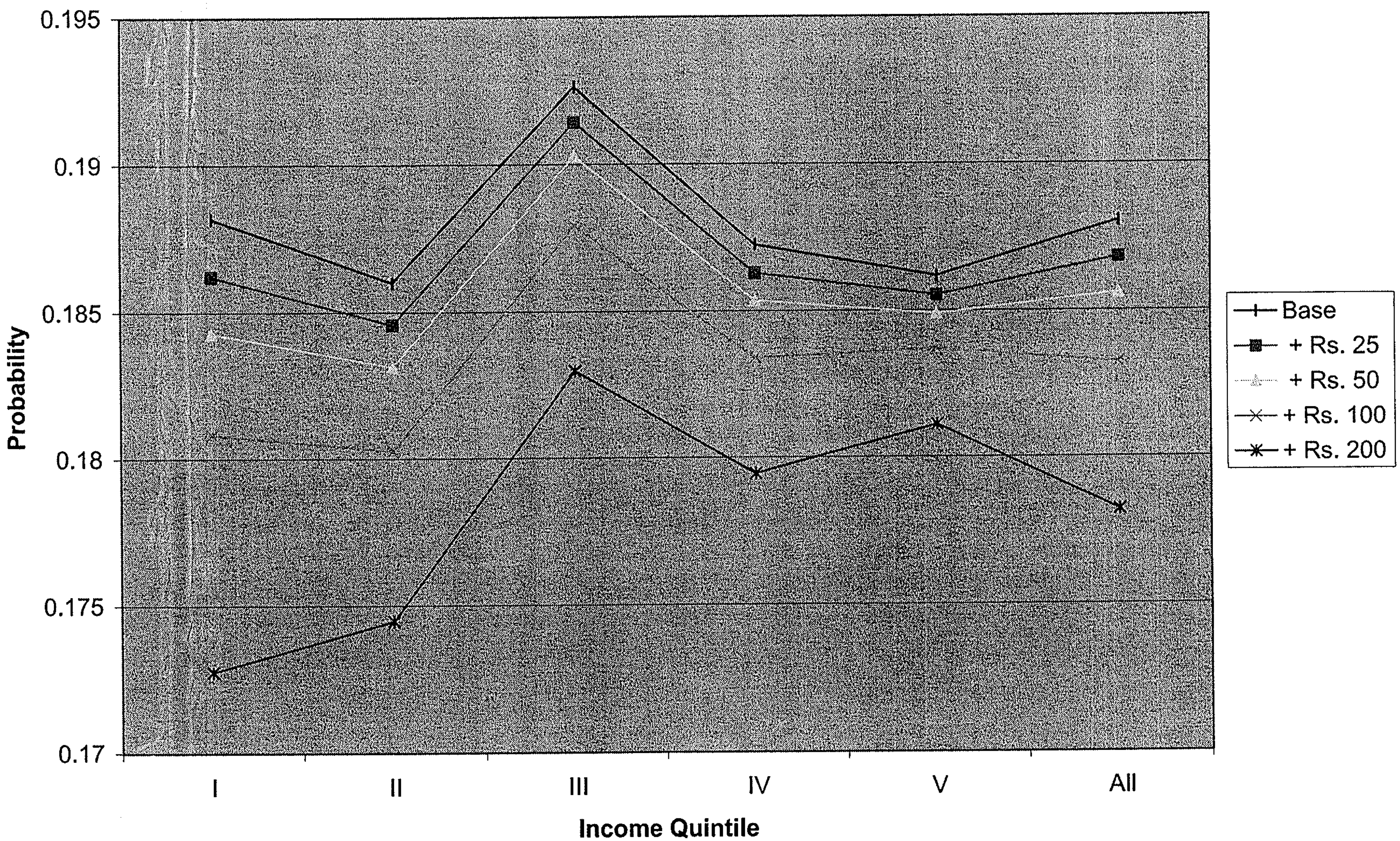
Table 9
FIML Estimation Results of Health Care Demand

Variable	MNL	UMNL (RU2)
Government Facility		
Constant	0.6094 (3.795)	0.431 (2.337)
Log Consumption[a]	-0.7679 (-3.599)	-0.8077 (-3.458)
Log Cons. Squared[a]	0.088 (4.962)	0.0945 (4.894)
Distance a[a]	-0.2623 (-7.006)	-0.293 (-6.965)
Distance b[a]	-0.4959 (-12.864)	-0.555 (-11.897)
Distance c[a]	-0.7516 (-20.055)	-0.8375 (-16.759)
Treatments	-0.033 (-0.751)	-0.0386 (-0.952)
Age	0.00434 (3.035)	0.0053 (3.333)
AgeD	-0.00037 (-6.91)	-0.00004 (-6.401)
Female	-0.1705 (-2.936)	-0.162 (-2.611)
Education	0.1902 (2.20)	0.1864 (2.043)
Household Size	0.0378 (4.094)	0.036 (3.521)
SCST	-0.0344 (-0.600)	-0.0086 (-0.138)
Habits	-0.1685 (-3.276)	-0.016 (-2.950)
Drinking Water	-0.3392 (-3.793)	-0.2718 (-2.777)
Latrine	-0.1517 (-1.206)	-0.1294 (-0.963)
Private Facility		
Constant	1.1478 (7.878)	1.0474 (6.621)
Log Consumption[a]	-0.7679 (-3.599)	-0.8077 (-3.458)
Log Cons. Squared[a]	0.088 (4.962)	0.0945 (4.894)
Distance a[a]	-0.2623 (-7.006)	-0.293 (-6.965)
Distance b[a]	-0.4959 (-12.864)	-0.555 (-11.897)
Distance c[a]	-0.7516 (-20.055)	-0.8375 (-16.759)
Treatments	0.1266 (3.723)	0.1415 (4.034)
Age	0.00299 (2.175)	0.0037 (2.499)
AgeD	-0.00035 (-6.875)	-0.00036 (-6.561)
Female	-0.0258 (-4.603)	-0.2634 (-4.470)
Education	0.2103 (2.547)	0.2154 (2.463)
Household Size	0.1762 (1.961)	0.0131 (1.377)
SCST	-0.1916 (-3.375)	-0.1922 (-3.171)
Habits	-0.1588 (-3.174)	-0.1528 (-2.873)
Drinking Water	-0.7709 (-7.597)	-0.7717 (-7.111)
Latrine	-0.5131 (-4.484)	-0.5388 (-4.472)
Private Doctor		
Constant	1.6034 (11.999)	1.5749 (11.474)
Log Consumption[a]	-0.7679 (-3.599)	-0.8077 (-3.458)
Log Cons. Squared[a]	0.088 (4.962)	0.0945 (4.894)
Distance a[a]	-0.2623 (-7.006)	-0.293 (-6.965)
Distance b[a]	-0.4959 (-12.864)	-0.555 (-11.897)
Distance c[a]	-0.7516 (-20.055)	-0.8375 (-16.759)
Treatments	-0.1708 (-4.270)	-0.1945 (-4.725)

Age	-0.0037 (-3.151)	-0.004 (-3.246)
AgeD	-0.0003 (-6.073)	-0.0003 (-5.834)
Female	-0.2403 (-4.979)	-0.2431 (-4.933)
Education	0.1094 (1.470)	0.0974 (1.282)
Household Size	0.066 (8.681)	0.0688 (8.741)
SCST	-0.1901 (-2.265)	-0.1006 (-2.019)
Habits	-0.2229 (-5.063)	-0.227 (-5.047)
Drinking Water	-0.8438 (-10.113)	-0.8609 (-9.832)
Latrine	-0.2977 (-2.804)	-0.2916 (-2.714)
Inclusive Value (RU2)		
Self		1.00 (fixed)
Formal		0.8677 (20.572)
Inclusive Value (RU3)		
Self		1.00 (fixed)
Formal		1.1525 (20.572)
Log-likelihood	-19157.50	-19153.88
Observations	16668	16668

[a] The coefficients of log consumption, log consumption squared, and all distance variables are restricted to be equal across alternatives.
Figures in parentheses are t-statistics.

Figure 3a: Mean Predicted Probabilities by Income Quintile (Gov Facility)



Foigure 3b: Mean Predicted Probabilities by Income Quintile (Pvt Facility)

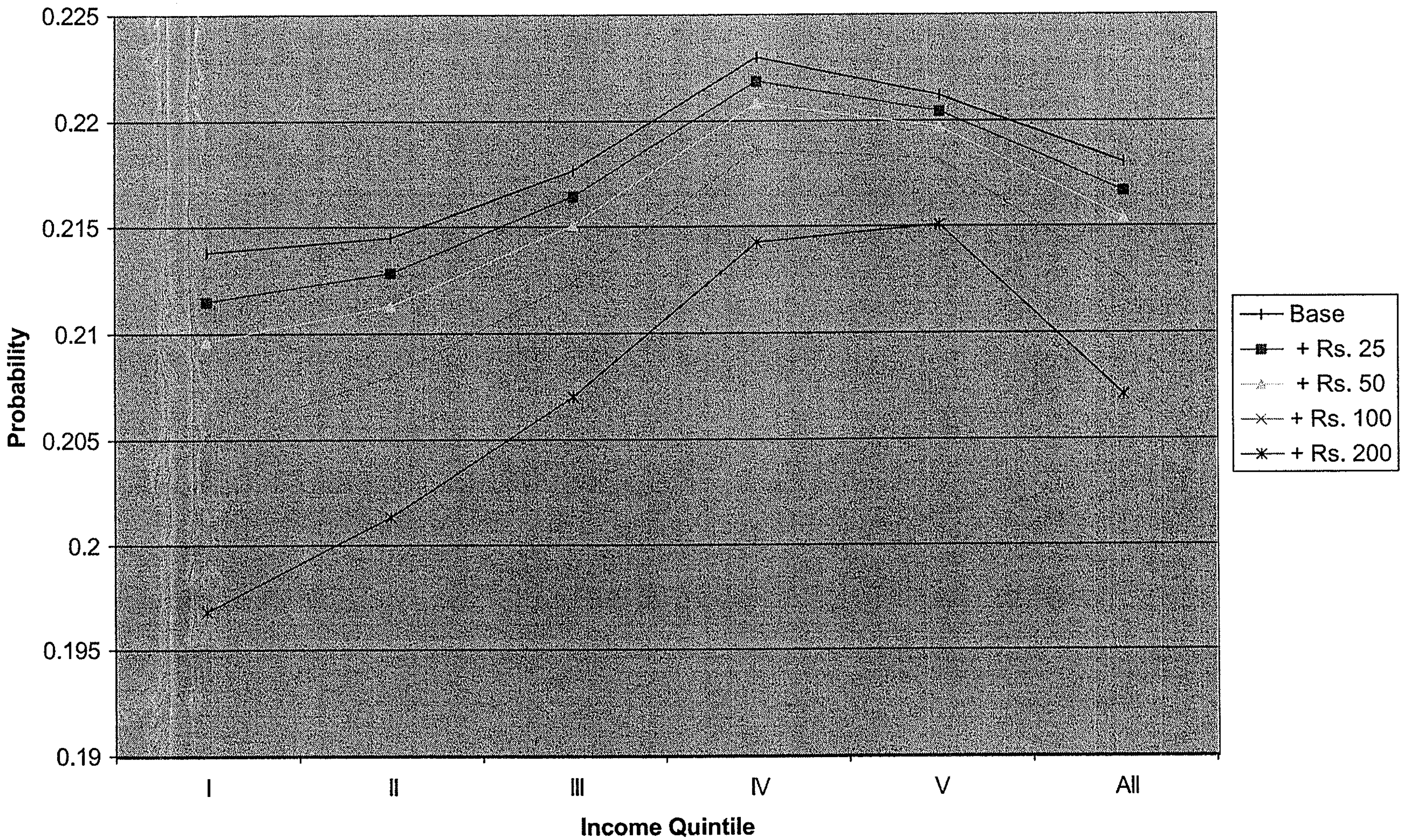
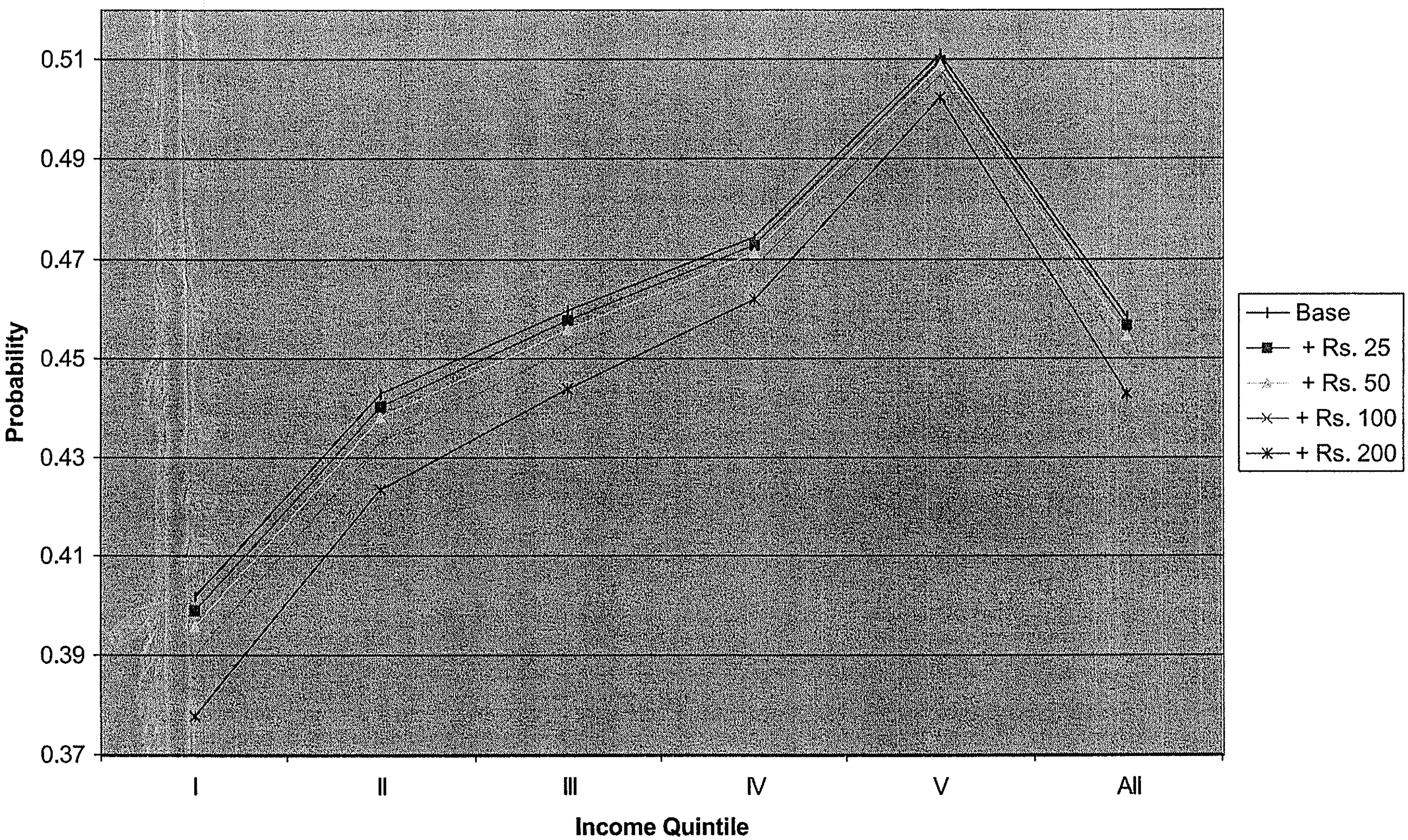


Figure 3c: Mean Predicted probabiliies by Income Quintile (Doctor's Clinic)



Simulation Results

Table 10 (a): Gov Facility: Total Sample

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	2754.68	16.53	2731.419	16.39	-0.136	2708.67	16.26	-0.269	2663.05	16.00	-0.525	2561.680	15.46	-1.069
Pvt Facility	3276.58	19.66	3281.206	19.69	0.032	3285.605	19.72	0.062	3291.255	19.78	0.119	3299.55	19.91	0.252
Pvt Doctor	7260.74	43.56	7269.777	43.63	0.065	7278.743	43.69	0.126	7292.453	43.82	0.259	7314.446	44.14	0.576
Self Care	3376.00	20.25	3381.597	20.29	0.038	3387.98	20.33	0.080	3395.24	20.40	0.147	3396.327	20.49	0.240
Total	**16668.00**	100.00	**16664.00**	100.00	0.000	**16661.00**	100.00	0.000	**16642.00**	100.00	0.000	**16572.00**	100.10	0.103

Gov Facility: Quintile 1

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	488.6197	15.03	481.4861	14.82	-0.201	474.9135	14.63	-0.394	462.53	14.30	-0.723	428.386	13.49	-1.533
Pvt Facility	599.0999	18.42	599.6172	18.46	0.039	600.6067	18.50	0.080	600.379	18.56	0.142	595.13	18.74	0.322
Pvt Doctor	1167.621	35.90	1169.067	35.99	0.089	1170.527	36.06	0.156	1171.254	36.22	0.312	1165.982	36.72	0.819
Self Care	996.6601	30.65	997.83	30.72	0.074	999.95	30.81	0.158	999.84	30.92	0.269	985.503	31.04	0.392
Total	**3252.00**	100.00	**3248.00**	100.00	0.000	**3246.00**	100.00	0.000	**3234.00**	100.00	0.000	**3175.00**	100.00	0.000

Gov Facility: Quintile 2

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	550.6482	16.54	545.2433	16.38	-0.162	539.8248	16.22	-0.325	528.8857	15.89	-0.649	506.24	15.24	-1.297
Pvt Facility	619.7568	18.62	620.9062	18.65	0.035	622.0579	18.69	0.069	624.101	18.75	0.136	627.66	18.90	0.283
Pvt Doctor	1389.617	41.74	1392.139	41.82	0.076	1394.669	41.89	0.152	1399.39	42.05	0.306	1406.050	42.34	0.595
Self Care	768.978	23.10	770.7109	23.15	0.052	772.45	23.20	0.104	775.6207	23.31	0.207	781.0499	23.52	0.419
Total	**3329.00**	100.00	**3329.00**	100.00	0.000	**3329.00**	100.00	0.000	**3328.00**	100.00	0.000	**3321.00**	100.00	0.000

Gov Facility: Quintile 3

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	576.1166	17.28	571.5645	17.14	-0.137	567.1617	17.01	-0.269	557.62	16.74	-0.540	538.61	16.18	-1.101
Pvt Facility	633.4888	19.00	634.5173	19.03	0.031	635.5227	19.06	0.061	637.0489	19.12	0.124	640.9742	19.25	0.253
Pvt Doctor	1456.107	43.67	1458.283	43.74	0.065	1460.400	43.80	0.129	1463.389	43.93	0.258	1471.663	44.21	0.533
Self Care	668.2878	20.04	669.6359	20.09	0.040	670.92	20.12	0.079	672.94	20.20	0.158	677.7524	20.36	0.314
Total	**3334.00**	100.00	**3334.00**	100.00	0.000	**3334.00**	100.00	0.000	**3331.00**	100.00	0.000	**3329.00**	100.00	0.000

Gov Facility: Quintile 4

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	571.8777	16.80	568.14	16.69	-0.110	564.26	16.58	-0.219	556.53	16.36	-0.441	541.36	15.91	-0.887
Pvt Facility	694.2387	20.39	695.1551	20.42	0.027	695.7278	20.44	0.050	697.358	20.50	0.104	701.1063	20.61	0.214
Pvt Doctor	1570.493	46.14	1572.387	46.19	0.056	1573.864	46.25	0.113	1577.302	46.36	0.227	1584.960	46.59	0.452
Self Care	567.3906	16.67	568.3184	16.70	0.027	569.15	16.72	0.057	570.81	16.78	0.110	574.572	16.89	0.221
Total	**3404.00**	100.00	**3404.00**	100.00	0.000	**3403.00**	100.00	0.000	**3402.00**	100.00	0.000	**3402.00**	100.00	0.000

Gov Facility: Quintile 5

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	567.4916	16.95	564.9854	16.87	-0.075	562.5105	16.80	-0.149	557.4878	16.66	-0.289	547.0845	16.36	-0.590
Pvt Facility	730.3174	21.81	731.0103	21.83	0.021	731.6894	21.85	0.041	732.3687	21.88	0.074	734.6741	21.96	0.156
Pvt Doctor	1676.508	50.06	1677.902	50.10	0.042	1679.283	50.14	0.083	1681.114	50.23	0.168	1685.791	50.40	0.337
Self Care	374.6838	11.19	375.1016	11.20	0.012	375.518	11.21	0.025	376.0296	11.23	0.047	377.4498	11.28	0.096
Total	**3349.00**	100.00	**3349.00**	100.00	0.000	**3349.00**	100.00	0.000	**3347.00**	100.00	0.000	**3345.00**	100.00	0.000

Simulation Results

Table 10 (b) Pvt Facility: Total Sample

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Pvt 20%	% change	+ Rs. 200	Percent	% change
Gov Facility	2754.68	16.53	2757.75	16.56	0.036	2760.604	16.59	0.064	2764.17	16.67	0.139	2773.14	16.80	0.275
Pvt Facility	3276.58	19.66	3246.448	19.50	-0.160	3222.676	19.37	-0.290	3158.791	19.04	-0.613	3049.171	18.47	-1.184
Pvt Doctor	7260.74	43.56	7267.576	43.65	0.088	7275.040	43.72	0.162	7284.789	43.92	0.360	7306.516	44.27	0.708
Self Care	3376.00	20.25	3378.23	20.29	0.035	3380.68	20.32	0.063	3378.25	20.37	0.114	3376.177	20.46	0.201
Total	**16668.00**	100.00	**16650.00**	100.07	0.000	**16639.00**	100.00	0.000	**16586.00**	100.00	0.000	**16505.00**	100.00	0.000

Pvt Facility: Quintile 1

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Pvt 20%	% change	+ Rs. 200	Percent	% change
Gov Facility	488.6197	15.03	488.09	15.08	0.053	487.73	15.10	0.075	484.93	15.22	0.191	480.08	15.37	0.347
Pvt Facility	599.0999	18.42	587.741	18.16	-0.266	582.8593	18.05	-0.377	559.3547	17.55	-0.871	527.8413	16.90	-1.521
Pvt Doctor	1167.621	35.90	1166.369	36.03	0.128	1166.148	36.10	0.199	1159.184	36.37	0.468	1149.390	36.80	0.899
Self Care	996.6601	30.65	994.80	30.73	0.084	993.27	30.75	0.104	983.54	30.86	0.213	965.6878	30.92	0.274
Total	**3252.00**	100.00	**3237.00**	100.00	0.000	**3230.00**	100.00	0.000	**3187.00**	100.00	0.000	**3123.00**	100.00	0.000

Pvt Facility: Quintile 2

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Pvt 20%	% change	+ Rs. 200	Percent	% change
Gov Facility	550.6482	16.54	551.80	16.59	0.045	552.80	16.61	0.070	554.38	16.68	0.137	557.55	16.81	0.268
Pvt Facility	619.7568	18.62	613.9856	18.45	-0.162	608.1406	18.27	-0.343	595.2956	17.91	-0.708	570.08	17.19	-1.430
Pvt Doctor	1389.617	41.74	1392.356	41.85	0.107	1394.340	41.90	0.154	1398.716	42.08	0.337	1407.597	42.44	0.693
Self Care	768.978	23.10	770.86	23.17	0.070	772.72	23.22	0.119	775.61	23.33	0.234	781.7731	23.57	0.469
Total	**3329.00**	100.00	**3329.00**	100.06	0.060	**3328.00**	100.00	0.000	**3324.00**	100.00	0.000	**3317.00**	100.00	0.000

Pvt Facility: Quintile 3

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Pvt 20%	% change	+ Rs. 200	Percent	% change
Gov Facility	576.1166	17.28	576.91	17.31	0.034	577.8236	17.35	0.067	579.63	17.42	0.142	583.48	17.55	0.273
Pvt Facility	633.4888	19.00	628.2318	18.85	-0.146	623.316	18.71	-0.288	611.956	18.39	-0.607	591.6131	17.80	-1.203
Pvt Doctor	1456.107	43.67	1457.740	43.75	0.075	1459.67	43.82	0.146	1463.691	43.99	0.320	1472.777	44.31	0.633
Self Care	668.2878	20.04	669.12	20.08	0.037	670.1892	20.12	0.075	671.72	20.19	**0.145**	676.1326	20.34	0.296
Total	**3334.00**	100.00	**3332.00**	100.00	0.000	**3331.00**	100.00	0.000	**3327.00**	100.00	0.000	**3324.00**	100.00	0.000

Pvt Facility: Quintile 4

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Pvt 20%	% change	+ Rs. 200	Percent	% change
Gov Facility	571.8777	16.80	571.8777	16.80	0.000	573.37	16.85	0.054	574.82	16.91	0.106	578.55	17.05	0.251
Pvt Facility	694.2387	20.39	694.2387	20.39	0.000	685.0098	20.14	-0.259	675.4996	19.87	-0.527	656.3086	19.34	-1.052
Pvt Doctor	1570.493	46.14	1570.493	46.14	0.000	1574.707	46.29	0.151	1578.973	46.44	0.304	1584.404	46.70	0.560
Self Care	567.3906	16.67	567.3906	16.67	0.000	568.91	16.72	0.054	570.71	16.79	0.117	573.7404	16.91	0.241
Total	**3404.00**	100.00	**3404.00**	100.00	0.000	**3402.00**	100.00	0.000	**3400.00**	100.00	0.000	**3393.00**	100.00	0.000

Pvt Facility: Quintile 5

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Pvt 20%	% change	+ Rs. 200	Percent	% change
Gov Facility	567.4916	16.95	567.4916	16.95	0.005	568.8801	16.99	0.047	570.4104	17.04	0.092	573.4871	17.22	0.277
Pvt Facility	730.3174	21.81	730.3174	21.81	0.007	723.3499	21.61	-0.202	716.685	21.41	-0.401	703.3236	21.12	-0.686
Pvt Doctor	1676.508	50.06	1676.508	50.07	0.015	1680.173	50.18	0.124	1684.225	50.31	0.245	1692.347	50.82	0.761
Self Care	374.6838	11.19	374.6838	11.19	0.003	375.5965	11.22	0.031	376.6801	11.25	0.063	378.8429	11.38	0.189
Total	**3349.00**	100.00	**3349.00**	100.03	0.030	**3348.00**	100.00	0.000	**3348.00**	100.00	0.000	**3348.00**	100.54	0.541

Simulation Results

Table 10 (C) Doctor: Total Sample

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	2754.68	16.53	2762.42	16.59	0.060	2771.36	16.65	0.125	2787.433	16.77	0.243	2814.98	17.02	0.490
Pvt Facility	3276.58	19.66	3285.41	19.73	0.070	3294.96	19.80	0.140	3313.60	19.93	0.277	3345.456	20.22	0.566
Pvt Doctor	7260.74	43.56	7219.727	43.35	-0.210	7177.964	43.13	-0.432	7099.795	42.71	-0.848	6926.886	41.87	-1.686
Self Care	3376.00	20.25	3386.45	20.33	0.080	3398.72	20.42	0.167	3421.180	20.58	0.328	3454.678	20.88	0.630
Total	**16668.00**	100.00	**16654.00**	100.00	0.000	**16643.00**	100.00	0.000	**16622.00**	100.00	0.000	**16542.00**	100.00	0.000

Doctor: Quintile 1

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	488.6197	15.03	489.5992	15.11	0.081	491.19	15.20	0.173	493.34	15.33	0.301	492.54	15.61	0.586
Pvt Facility	599.0999	18.42	599.50	18.50	0.075	600.342	18.57	0.152	602.5133	18.72	0.295	600.5108	19.03	0.611
Pvt Doctor	1167.621	35.90	1153.981	35.61	-0.299	1139.022	35.24	-0.663	1116.754	34.69	-1.212	1057.393	33.51	-2.390
Self Care	996.6601	30.65	997.917	30.79	0.143	1001.44	30.99	0.338	1006.39	31.26	0.616	1004.552	31.84	1.192
Total	**3252.00**	100.00	**3241.00**	100.00	0.000	**3232.00**	100.00	0.000	**3219.00**	100.00	0.000	**3155.00**	100.00	0.000

Doctor: Quintile 2

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	550.6482	16.54	552.65	16.61	0.070	555.03	16.68	0.137	59.67	1.79	-14.746	567.05	17.13	0.585
Pvt Facility	619.7568	18.62	622.03	18.70	0.080	624.6256	18.77	0.152	629.7611	18.94	0.323	637.9195	19.27	0.650
Pvt Doctor	1389.617	41.74	1380.657	41.50	-0.244	1372.168	41.23	-0.512	1353.081	40.69	-1.049	1312.298	39.63	-2.108
Self Care	768.978	23.10	772.66	23.22	0.124	776.18	23.32	0.223	782.48	23.53	0.434	793.7311	23.97	0.873
Total	**3329.00**	100.00	**3328.00**	100.03	0.030	**3328.00**	100.00	0.000	**2825.00**	84.96	-15.038	**3311.00**	100.00	0.000

Doctor: Quintile 3

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	576.1166	17.28	577.76	17.34	0.060	579.79	17.41	0.126	583.46	17.53	0.252	591.86	17.79	0.510
Pvt Facility	633.4888	19.00	635.429	19.07	0.070	637.7596	19.15	0.145	641.8035	19.28	0.284	651.4243	19.58	0.579
Pvt Doctor	1456.107	43.67	1448.059	43.46	-0.215	1440.392	43.24	-0.432	1424.635	42.81	-0.867	1394.189	41.91	-1.769
Self Care	668.2878	20.04	670.75	20.13	0.086	673.06	20.21	0.161	678.10	20.38	0.331	689.5274	20.73	0.681
Total	**3334.00**	100.00	**3332.00**	100.00	0.000	**3331.00**	100.00	0.000	**3328.00**	100.00	0.000	**3327.00**	100.00	0.000

Doctor: Quintile 4

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	571.8777	16.80	573.60	16.85	0.050	575.23	16.90	0.103	578.51	17.01	0.205	585.65	17.31	0.511
Pvt Facility	694.2387	20.39	696.3838	20.46	0.063	698.4294	20.52	0.129	702.597	20.65	0.258	711.5023	21.03	0.637
Pvt Doctor	1570.493	46.14	1564.663	45.97	-0.171	1558.157	45.79	-0.349	1545.730	45.44	-0.701	1520.494	44.95	-1.192
Self Care	567.3906	16.67	569.36	16.73	0.058	571.18	16.78	0.116	575.16	16.91	0.238	583.3564	17.24	0.575
Total	**3404.00**	100.00	**3404.00**	100.00	0.000	**3403.00**	100.00	0.000	**3402.00**	100.00	0.000	**3401.00**	100.53	0.532

Doctor: Quintile 5

Choice	Base	Percent	+ Rs. 25	Percent	% change	+ Rs. 50	Percent	% change	+ Rs. 100	Percent	% change	+ Rs. 200	Percent	% change
Gov Facility	567.4916	16.95	568.8102	16.98	0.039	570.1261	17.02	0.079	572.4424	17.10	0.153	577.8773	17.26	0.315
Pvt Facility	730.3174	21.81	732.0587	21.86	0.052	733.7983	21.91	0.104	736.9203	22.01	0.204	744.099	22.23	0.418
Pvt Doctor	1676.508	50.06	1672.367	49.94	-0.124	1668.225	49.81	-0.247	1659.593	49.57	-0.490	1642.512	49.06	-1.000
Self Care	374.6838	11.19	375.7624	11.22	0.032	376.8499	11.25	0.065	379.0457	11.32	0.134	383.5109	11.45	0.267
Total	**3349.00**	100.00	**3349.00**	100.00	0.000	**3349.00**	100.00	0.000	**3348.00**	100.00	0.000	**3348.00**	100.00	0.000

Determinants of Elderly Living Arrangements: Evidence from Aging in Manitoba Longitudinal Data, 1971-96

Abstract

Linking Aging In Manitoba (AIM) longitudinal study on 1971 cohort's interview data with home care admission data, this study analyzes the determinants of elderly living arrangements. It is found that home care utilization reduces the demand for nursing home and increases the demand for independent living. Loss of a spouse affects independent living negatively and both cohabiting and nursing home residence positively. The effect of age on nursing home residence is positive and on independent living and cohabiting is negative. Educated people are more likely to live independently than cohabit or enter an institution. Similarly, those who are healthy and satisfied in life are more likely to live independently instead of cohabiting or entering nursing homes. Those who lived longer in the community are more likely to live independently or cohabit rather than enter an institution. Home ownership is positively associated with both independent living and cohabiting. The results are suggestive of possible income related inequity in institutionalization.

Key Words: long-term care, elderly, living arrangement, Canada

JEL Codes: J14, I18, I11

1 Introduction and the Literature

It is well known that Canada's population is growing older, due to aging of the population as well as declining fertility and mortality. According to Statistics Canada, the proportion of those aged 65 and over in Canada was 13% in 2000, comprising 4 million people, and is expected to increase to 21% by 2026, to about 8 million people. The number of Canadians over the age of 65 will more than double, from 3.9 million in 2000 to 9.3 million in 2040. However, the most rapidly growing age group in Canada will be 80 and older, projected to increase from 920,000 in 2000 to 1.9 million in 2026. The demographic composition of Manitoba exhibits a similar trend.

There can be several serious economic consequences of this demographic challenge. Aging of the population, coupled with onset of disability, loss of a spouse, and deterioration in health status, may lead to: 1) a higher demand for old age income security; 2) an increase in demand for health, medical, and personal care; and 3) revised decisions in living arrangements. Recent research in the United Kingdom shows that very old people are more likely to have a long-standing illness, to be more dependent, to have more functional problems, and to have received more health care Tinker et al. [61]. Studies in Canada also suggest that aging of the population would lead to a growing demand for health, medical, and long-term care, among other things (Denton and Spencer [14], [15], [16], [17], and Rosenberg [57]). Various reports at Manitoba Centre for Health Policy (MCHP) show that elderly people tend to use more health care, including home care, hospitals, prescription drugs and nursing homes in the Province of Manitoba, Canada. The contribution of elderly morbidity to the higher use of physician services is also well documented in Manitoba, Black et al. [5]. Further, older women are more likely to be widowed, live alone, have poor health and be worse off financially, Tinker et al. [61].

Although elderly Canadians are generally healthy, there is uncertainty about demand for certain types of health services, including personal care. Uncertainty arises not only due to incidence of illness, disability, and non-medical life cycle events (such as loss of a spouse) but the provision of different types of care is indeterminate. In this study, we intend to explore the underlying factors governing the choice of living arrangements of the elderly and subsequently discuss the long-term care issues.

Studies in the United States show that the demand for long-term care by the elderly are closely associated with the choice of living arrangements (Börsch-Supan [6], Börsch-Supan et al.[7], Kotlikoff and Morris [40], Stern

[60], and Hoerger et al. [32]). The theoretical motivations of these studies in the United States are quite interesting, however, they serve little purpose for public policy in Canada due to the distinct institutional set up and the nature of public intervention. There has been a rapid expansion of home care programs, the effects of which on independent living, cohabiting and institutionalization are not well documented in the literature. More recently, Roos et al. [54] reported that between 1990 and 1997, home care spending in Manitoba as well as in all other provinces more than doubled. Thus, from a public policy point of view, important questions remain unanswered: Do home care programs affect living arrangement decisions? Does the government need to spend more resources on home care programs or nursing home programs? The choice of living arrangements by the elderly may also be influenced by various other needs, economic conditions, home ownership, etc.

There are a number of studies, however, relating to home care utilization and nursing home entry in Canada (Tomiak et al. [62], Tomiak et al. [63], Carriere and Pelletier [8], Trottier et al. [64], and Roos et al. [54]). Although some of these studies reveal some information on long-term care use, their analysis is confined to cross-sectional and cross-tabulation data analysis with an exception of Tomiak et al. [63]). Using the 1986 census data linked to Manitoba longitudinal health care utilization data, Tomiak et al. [63] discuss the effect of living arrangements on nursing home entry in Manitoba. By adjusting for all other variables from their model, they found that presence of the spouse does not reduce the risk of nursing home entry, whilst having an additional household member does reduce the hazard of nursing home entry. However, many important variables like marital status, widowhood, income, etc. are confined to the 1986 census and do not vary with the health care utilization data. A number of other studies, on the other hand, show that widowhood, income, and other non-medical life-cycle events are important factors in explaining the choice of living arrangements (Börsch-Supan [6], Börsch-Supan et al. [7], and Hoerger et al. [32]).

A study of elders in Saskatchewan suggests that persons receiving light home care (such as homemaking, personal care, meal provision and other measures designed to help them live in the community) are 50% more likely to die and 50% more likely to enter a health institution than those receiving nursing services (Health Services Utilization and Research Commission [30]). A study from British Columbia suggests that the cost of formal home care was less expensive than nursing home care, Hollander [33]. A study in the United

States also found that home care was more cost-effective than the nursing home care, Kane et al. [38]). A report by MCHP on home care use states that 93% of persons admitted to a nursing home were home care clients, and the average number of days of home care received was 537, Roos et al. [55]. Thus, the effect of the provision of formal home care on institutionalization is unclear. In this study, we take a more unifying approach to examine the determinants of living arrangements, including the effects of home care utilization, where institutionalization is an option.

The institutional living arrangement is a choice in the sense that an individual first made an application for moving into an institution for her/his long-term care needs. Thus, from a public policy point of view, understanding the underlying determinants of these forms of living arrangement decisions would assist the policy makers in allocating scarce resources among competing needs and address the issue of the provision of long-term care for the elderly. Specifically, we seek to answer the following set of questions:

- What are the effects of home care on living arrangement decisions? Does it help the seniors to stay independently or cohabit? Does it reduce the demand for nursing home?

- What are the important non-medical life-cycle events that are responsible for leaving the community and entering a health institution?

- What are the socio-demographic factors that play an important role in switching away from non-institutionalized living arrangements towards institutionalized living arrangements?

- Do health status and life satisfaction affect living arrangement decisions?

The rest of the paper is organized as follows. Section 2 briefly discusses the long-term care issues in Manitoba and empirical findings in the literature. In section 3, we briefly describe the Aging In Manitoba (AIM) data source and variable constructions for this study. We then illustrate our framework and empirical approach with a careful analysis of the micro data in section 4. Section 5 discusses our conclusions and the direction of future research on the economics of aging.

2 Long-term Care

2.1 Long-term Care Provisions

The insured nursing home program (officially known as personal care homes) in Manitoba started on July 1, 1973 (Management Committee of Cabinet [42]). Nursing home represents a form of living arrangement associated with a provision of care for those who can no longer be cared for at home. The services that are typically provided at a nursing home facility are basic nursing care, drugs and supervision of the activities of daily living. Although the eligibility to enter a nursing home is determined on the basis of need, residents are charged on a daily basis based on their income. Since August 1, 2000 the minimum rate of $26.30 per day and the maximum of $61.40 are charged. It has been found that the rate of admission to nursing homes is higher among those aged 85 and over. The age-sex adjusted rates of nursing home use have declined between 1985-86 and 1998-99, Roos et al. [56]. The waiting time for admission to nursing homes and their average length of stay have declined over the years, Menec et al. [50].

The Home Care Program in Manitoba started in 1974 (Management Committee of Cabinet [42]) and has expanded over the years. The mandate of the home care program has been to provide home care services to those who have inadequate informal sources of care (informal care is typically provided by the spouse and children) to return home from hospital or live independently in the community or while waiting to enter a nursing home. The regional health authorities are required to provide home care services free of charge to those who meet the need criteria of admission into the home care program. The home care program in Manitoba provides a range of services including nursing services, personal care assistance, palliative care, meal preparation, cleaning and laundry services, medical supplies, etc.

It has been reported that between 1990 and 1997 Manitoba experienced a 34% increase in home care users and 119% increase in expenditures on home care, Roos et al. [54]. It is also found that the days open to home care (an overall indicator of home care use defined both by the number of people who receive home care and their duration of service use) is on the rise increasing by 25% for the aged 65 to 74, 24% for the aged 75 to 84 and 17% for the 85+ year olds, Menec et al. [50]. The expenditures on home care during 1999-2000 were $149 million dollar per year, Manitoba Health [43]. A similar growth in home care programs is seen across Canada as well and is expected to expand due to demographic changes in the population and a shift from institutionalized care towards home care (Di Matteo and Di Matteo [13], and CIHI [10]).

2.2 Long-term Care Issues

A variety of factors have been investigated in the literature on the need for nursing home entry. A number of socio-demographic factors (such as age, gender, location, socio-economic status, etc.) have been associated with increased nursing home use. Higher age has been typically associated with higher admission rates in most studies. Marital status has been found to be an important predictor of the need for nursing home care. Married people are less likely to be institutionalized than widows and widowers, Mustard et al. [51].

Low income and low education in Manitoba are associated with a higher probability of nursing home entry (Mustard et al. [51] and Tomiak et. al. [63]). Home ownership, an important indicator of assets and accumulated/inherited wealth, is found to reduce institutionalization in England and Wales, Grundy and Glaser [23]. Many studies explore the effects of health status on nursing home admission. Various specific health conditions (such as diabetes, stroke, neurological disorders, musculoskeletal disorders, cancer, heart disease, dementia, Alzheimer's disease, etc.) have been found to be associated with higher nursing home use; the effects are ambiguous across studies though. Thus, self-perceived health status is used as a more general measure of health in many studies. Low health status is found to be associated with a higher likelihood of institutionalization, Steinbach [59].

There are conflicting findings with regard to gender though. Some studies found that female are more likely to be institutionalized (Rockwood et al. [53], and Lavery et al. [41]) than men, whereas others report the opposite (Freedman [20], Mustard et al. [51], and Smith et al. [58]). Typically, women live longer than men. As women are more likely to be widowed and lack informal sources of care, this may not be a good predictor of nursing home entry.

In this paper, we view the demand for long-term care as closely associated with individuals' living arrangement decisions. Retired individuals may prefer to live independently, or stay in an intergenerational family, or in a health institution. Of course, in the 'single-entry' system for long-term care in the province of Manitoba, individuals' care needs are assessed upon entry and services are provided according to their needs.

3 Data

3.1 Data Source

Aging in Manitoba (AIM) is the largest and longest longitudinal study on aging in Canada. Representative samples of elderly individuals living in Manitoba were interviewed in 1971, 1976 and 1983, respectively. Survivors from the 1971 and 1976 samples were interviewed again in 1983. In 1990, survivors of the three cohorts were re-interviewed and in 1996 and 2001 survivors were again interviewed. AIM interview data contain a wealth of socio economic information about these representative samples of elderly in each of their cross-sectional and panel data sets.

A random sample of 4,803 individuals, stratified by age and gender using a small area probability sampling frame of both community and institutional dwelling Manitobans aged 65 and over, were interviewed in 1971. The second interview was conducted in 1983 covered 1,518 survivors. The third and fourth interviews were conducted in 1990 and 1996 covered 630 and 214 survivors, respectively. In this paper, we use the 1971 cohort's interview data and the corresponding home care admission data from the administrative data base developed by Manitoba Health over 30 years. This unique linked data set, is the most appropriate for this study.

3.2 Data Description

The aggregate data on living arrangements shows that independent living has been the most dominant form of living arrangements among the elderly Manitobans. Table A.1 shows that more than 60% of the elderly lived independently between 1971 and 1990. Since 1971, the percentage of cohabiting (i.e., living in an intergenerational family) has been on the decline and the percentage of people entering a nursing home has been rising significantly. This suggests that aging of the elderly is one of the indicators of institutionalization. The age distribution of the elderly sample in 1971 is presented in Table A.2. It is clear to see from Table A.2 that the relatively young-old constitutes a very large proportion of the sample; the percentage of the people 85+ years old is around 11 percent. So, as the fraction of the survivors gets into the age cohort of 85+ years, institutionalization becomes the preferred form of living arrangement. The educational status of respondents at the 1971 baseline is presented in Table A.3. The level of education of the sample of elderly in 1971 is relatively lower than that of today's elderly people.

One of the most important non-medical life cycle events is loss of a spouse.

Widows and widowers are most likely to need formal and informal care and support in the society and community then the married for their daily living. Table A.4 shows that over the period 1971-1996, the percentage of the married has been declining dramatically, from 50 percent in 1971 to 16% of the total survivors in 1996. The percentage of widows and widowers are rising significantly during the same time period, from 37% in 1971 to 70% in 1996.

Associated with aging may be deteriorating health status, functional impairments and a general disorientation and dissatisfaction towards life among the elderly. Self-reported measures of health status and self-reported measures of general satisfaction towards life have been consistently collected in each of the AIM survey data sets. Tables A.5 and A.6 present the self-reported health status and self-reported general life-satisfaction responses over the four waves, respectively. The percentage of elderly reporting excellent, good or fair health and excellent, good or fair life-satisfaction has been declining marginally. However, the percentage of the respondents in 'not applicable category' has increased rapidly. Not applicable cases are a very special category in which the AIM study used proxy respondents due to higher incidence of physical and cognitive impairments. Typically, most of them were living in nursing homes and unable to respond to the interviewer.[1]

Those who lived longer in the community may be better off in old age and survive longer because of informal sources of care and emotional support provided by the family members and friends in the community. Table A.7 shows that the percentage of people who survived in 1996 was actually higher for those who lived in the community for more than 25 years. This suggests that provision of emotional and social support available in the community might have helped elders to stay healthier and survive longer.

Financial independence and income security in old age may also be a contributing factor towards the elderly well-being. Table A.8 reports monthly income from all sources as reported by the individuals in the survey. Given that the average monthly income is higher for the survivors in 1996, this might be suggestive of possible effect of income and wealth on mortality and health status.[2] Another indicator of financial asset and wealth is home ownership. Table A.9 reports the proportion of people (or their spouse) who

[1] In the subsequent analysis, we did not drop these not applicable cases because it led to a serious problem of loosing most individuals staying in nursing homes. We treated these categories as a possible state of poor health status, their health status cannot possibly be excellent or good on any measure because they crossed the threshold level of poor health to fulfil the nursing home entry requirements. Inclusion of these responses might overestimate the influence of self-reported health on living arrangements.

[2] It is to be noted that the self reported income is an inadequate measure and many respondents reported zero income. Further, the monthly income data reported here are in current dollars as of the year of each interview.

own a house (with or without a mortgage). One caveat, however, with home ownership response is that those who entered a nursing home no longer own a home; that is the response of home ownership and nursing home residence are mutually exclusive. It might be possible that those who were eligible to enter a nursing home have either disposed of their home to finance long-term care needs (minimum daily charges at nursing homes) or transferred ownership to their children.

3.3 Variable Construction

In this sub-section, we describe our variable construction from the data sources to analyze the determinants of living arrangements of the elderly. Our dependent variable in this study takes three possible outcomes: a) independent living (living alone or living with spouse), b) cohabiting (living with children, siblings, friends, parents, or grand children) and c) living in a nursing home. The information on these three forms of living arrangements has been constructed from the questions on type of housing and family household information collected by the AIM survey. If a response on the type of housing is personal care home then the living arrangement decision is nursing home. A response is cohabiting if the respondent lives with at least one of the following members: sibling/in-law, child/in-law, friend/unrelated person, parent/ in-law, grand child/in-law and does not live in a nursing home. If the respondent lives only with a spouse or other forms of living arrangements (such as living alone in a house or self contained suite or senior citizen's house or board & room) then the respondent is said to have an independent form of living arrangement.[3] The descriptive statistics on these three forms of living arrangements are reported in Table 1.

The home care received variable has been constructed from the 1971 - 96 home care admission data compiled by Manitoba Health. Since the provision of formal home care in Manitoba came into existence in 1974, the concept of a formal home care program was not applicable in 1971. In 1983, if an individual admitted into the home care during 1974 - 1983 then we assign a value of one to home care received in 1983; otherwise it is assigned a value of zero. Similarly, in 1990 and 1996 if the respondent admitted into formal home care between 1983 – 1990 and 1990 –1996 we assign one to home care received in 1990 and 1996; otherwise it is assigned a value of zero. Table 1 presents the descriptive statistics on the home care received variable in all

[3]We recognize that some of the senior citizen residential apartments have provisions for formal and informal sources of care for the elderly (i.e., formal or informal on-site home care). These residences are not subject to the formal regulations and standards of nursing homes.

the years.

Gender is represented by a dummy variable (female = 1, male = 0). Marital status is characterized by a dummy variable (widows, separated and divorced = 1, zero otherwise), which indicates that married and singles are the reference category. Age and income are continuous variables. Home ownership takes a value of one if the respondent or spouse owns a home with or without mortgage, zero otherwise. Educational status of the respondent is characterized by two dummies for those respondents who completed 5 to 10 years of education and more than 10 years of education, leaving with less than 5 years of education as the reference category. Self-reported health is captured by dummy variables for excellent and good, leaving other responses as the reference category. Similarly, the general life satisfaction represented by two dummies for excellent life and good life, leaving fair, poor, bad and not applicable responses as the reference category. The number of years lived in the community is characterized by two dummies for those who lived more than 25 years and 11 – 25 years, leaving 10 years or less as the reference category. The descriptive statistics are presented in Table 1 for all four waves.

4 The Framework

Since the retired individuals are expected to take decisions on their living arrangements in an uncertain future environment, we assume that an individual or family choose a type of living arrangement so as to maximize the expected life-time utility. The framework is the random utility model, similar to Börsch-Supan [6], Börsch-Supan et al. [7], and Hoerger et al. [32] for modelling the elderly living arrangements. The elderly person can stay in the community with independent living arrangement or in an intergenerational family or move into a nursing home.[4] These three forms of the living arrangements are associated with different levels of formal care provisions.

Let a person's utility of choosing the living arrangement j ($j = 1, 2, 3$ representing independent, cohabiting and nursing home, respectively) be $U_j(F, C, \mathbf{X})$, where F is formal long-term care received, C is the consumption of other goods and services and $\mathbf{X}$ is a vector of person-specific characteristics. Assuming that the price of consumption is a numeraire, the price of formal care P_t and income Y_t at time period t, the elderly person's objective

[4]We ignore the issues of strategic interaction between elderly persons and adult children using game theoretic models (see Engers and Stern (2002) and the references cited in that paper). This literature may be relevant to explain why intergenerational living arrangement is on the decline everywhere.

is to choose F_t and C_t in order to maximize the expected lifetime utility for each living arrangement subject to the budget constraint. Formally, the dynamic optimization problem is:

$$MaxE_t \sum_{t=0}^{T} \delta \left[U\left(F_t, C_t, \mathbf{X}_t\right) \right]$$

subject to $C_t + P_t F_t + Y_t + S_t$, where δ is the discount factor and S_t is the level of subsidy from the provincial government. Basically, the government decides the levels of subsidy on the basis of need and income.

Let the indirect utility for state j in time period t be $V_{jt}(P, Y, \mathbf{X})$ and after solving the dynamic optimization problem an elderly person chooses alternative j if and only if $V_{jt}(P, Y, \mathbf{X}) > V_{kt}(P, Y, \mathbf{X}) \forall j \neq k$. However, moving into a nursing home is conditioned by a rationing rule. The expert panel screens all applicants and if the composite health condition of the individual exhibits a threshold limit then the person is allowed to enter a nursing home. Similarly, there is a rationing rule dealing with those who are entitled for free provision of home care. It is not possible to incorporate the rationing rules in the empirical analysis of this paper because we have no information on the screening process and their associated health conditions.[5] Disregarding

[5]Given the rich data source of Manitoba Health, it is possible to estimate the full

the rationing rule, one can apply a multiperiod multinomial probit or logit model to the living arrangement dynamics. In order to estimate a multiperiod multinomial choice model, one has to evaluate high dimensional integrals for the likelihood function. Analytical evaluation of such integrals is impossible and simulation based methods are generally employed to evaluate the likelihood function numerically.

Given the computational complexity of solving high dimensional integrals using numerical approximation, we use a relatively simple method in this paper.[6] We pool all the valid observations from the 1971 cohort of individuals and choose a simple multinomial logit model with year-specific fixed-effects.[7] The multinomial logit model for individual i choosing living arrangement

structural model incorporating the rationing rules for provision of home care and nursing home care. However, additional data from Manitoba Health on health care utilization data on AIM study participants are required.

[6]The problem is further complicated by the fact that many individuals fall into the absorbing state, i.e., death. Dealing with an absorbing state in a multinomial choice setting and introducing unobserved heterogeneity is computationally complex. Because of heavy computation involved, it takes more than a day to solve a random effects multinomial logit model for the reduced form model. Nonetheless, refinements involving panel estimation is our plans for future research work.

[7]One advantage of the multinomial logit functional form is its robustness against bias from self-selection (see McFadden, 1984).

type j is given by:

$$\Pr\left(V_{ij} > V_{ik}, \forall k \neq j\right) = \frac{\exp\left(\mathbf{X}_i^T \boldsymbol{\beta}_j\right)}{\sum_{k=1}^{3} \exp\left(\mathbf{X}_i^T \boldsymbol{\beta}_k\right)}; j = 1, 2, 3.$$

Where, $\mathbf{X}i$ is a vector of person specific characteristics and the year dummies. For model identification, we normalize the vector $\boldsymbol{\beta}_1$ to zero. That is, the expected utility from the reference alternative (i.e., independent living arrangement in this case) has been normalized to zero. So, the results are interpreted in relation to the independent living arrangement.

4.1 Multinomial Logit Estimates

The vector $\mathbf{X}_i$ in the multinomial logit model includes home care received, age, gender, marital status, education, income, home ownership, health status, general life satisfaction and the years lived in the community. However, a potential problem with inclusion of home care, home ownership, self-reported health status, general life satisfaction and the years lived in the community in a living arrangement choice framework is that they may be simultaneously determined causing endogenous bias. The appropriate procedure would be to find suitable instruments for these variables and apply instrumental variable method of estimation. However, finding suitable instruments is not an easy task; weak instruments are as problematic as endogenous regressors. Instead, we adopt a more direct approach of estimating a reduced form model without these covariates then add them to see if the results are sensitive.

Before we turn our discussion to the estimated results, it is necessary to present the results of the Independence of Irrelevance Alternatives (IIA) test. One crucial assumption of the multinomial logit model is that the ratio of the probabilities of the two alternatives j and k depend only on $\mathbf{X}_j$ and $\mathbf{X}_k$, and not on the presence of any other alternatives. This is known as IIA property. If the IIA property does not hold then the model is misspecified and the estimates are misleading, hence cannot be relied upon. For example, in the context of our study IIA implies that the ratio of the probabilities of independent living and cohabiting will not change in the presence of another alternative i.e., nursing home. Similar interpretation holds for any pair of alternatives.

In order to test the IIA property, a well known procedure has been developed by Hausman and McFadden [25]. The test procedure is to first estimate the multinomial logit model on a full set of alternatives, and second on a specified subset of alternatives. If the IIA property holds the two estimates should not be statistically significantly different. On the other hand, if the IIA property does not hold, there will be sharp differences within the subsets. And, the estimates from the second model will be larger in magnitude than the estimates from the full set of alternatives.

In the reduced form model, after dropping alternative 2 (i.e., cohabiting) and alternative 3 (i.e., nursing home), the Hausman test statistics are 18.946 and 11.448, respectively. With 11 degrees of freedom, we fail to reject the IIA. Thus, there is no evidence that the IIA property has been violated. Similarly, in model 2 of Table 2, the corresponding Hausman test statistics are 14.133 and –4.471. The chi square test of 14.902 with 19 degrees of freedom gives a p-value is 0.776, which fails to reject the IIA. For negative Chi-square test statistic, the model does not meet the asymptotic assumptions of the Hausman test. Hausman and McFadden (1984) recognized this possibility and conclude that a negative test result is evidence that IIA has not been violated. Thus, there is no evidence that the IIA property has been violated in this study.

We also checked whether or not any two alternatives should be combined using the Cramer-Ridder likelihood ratio test (Cramer and Ridder [11]). In the reduced form model, combining cohabiting and nursing home, cohabiting and independent and independent and nursing homes yield LR test statistics 265.9, 1136.74 and 798.85, respectively. The corresponding test statistics for model 2 are 306.84, 2880.8 and 2021.9, respectively. Since the Cramer-Ridder test results are quite significant across all possibilities, we cannot pool any pair of alternatives of living arrangements in this study.

The multinomial logit estimates for both specifications are presented in Table 2. Since the coefficients of independent living arrangement are constrained to be zero, the remaining coefficients are interpreted as relative to what they are for the independent living. The effect of age on the probability of living in nursing homes is positively significant and insignificant for cohabiting in both specifications. Gender was significant in the reduced form equation, but turned out to be insignificant when additional covariates were added.

Loss of a spouse increases the probability of cohabiting in both specifications, and its effect on the probability of choosing nursing homes is positively significant in the reduced form, but insignificant after adding other variables. The effect of income is positive on the demand for nursing home and negatively significant on the probability of cohabiting. In fact, for the cohabiting equation income appears to have a non-linear relationship, it is convex. The effect of education on both the probability of cohabiting and institutionalization is negatively significant in both specifications. This implies that educated people are more likely to choose an independent form of living arrangement. Although the magnitudes of education coefficients are reduced after adding additional regressors, the sign and statistical significance are preserved. This suggests that educated people are more likely to be active, perhaps because of their healthy life style and community participation at the old age, and live independently in the community.

The effect of home care admission (i.e., ex ante home care utilization) on the probability of cohabiting and choosing nursing home residence is negative. The coefficient in the cohabiting equation is relatively smaller and significant at the 5% level whereas the coefficient in the nursing home equation is much larger and significant at the 1% level. This suggests that provision of formal home care reduces demand for nursing home care and to a smaller extent cohabiting as well. The effect of home ownership is negative on the probability of choosing nursing home, which implies a homeowner is less likely to be institutionalized.

Those who are healthy and satisfied in life are more likely to remain independent and choose independent form of living arrangement. The health status and life satisfaction variables are negatively significant on the probability of cohabiting and choosing nursing home in relation to independent living. Those who lived longer in the community would choose to remain in the community and cohabit with at least one additional person. On the other hand, those who lived longer in the community are less likely to choose institutional form of living arrangement - the coefficients are negatively significant on the probability of choosing nursing home.

4.2 Individual Heterogeneity

The multinomial logit estimates of the previous section allowed year-specific fixed effects to account for institutional changes and the group characteristics of the survivors of 1971 cohort. However, unobserved individual heterogeneity is almost inevitable in micro data (Heckman, 2000), and cannot be easily accounted for in cross-sectional or pooled estimates. The random effects specification is appropriate when individual specific effects are uncorrelated with the included explanatory variables, if consistent estimates of the model parameters are of interest. Given the computational complexity of estimating multiperiod multinomial models discussed above, we estimate random effects logit model for each type of living arrangements separately to assess the effect of individual heterogeneity.

The random effects logit model for an unbalanced panel can be stated in terms of the standard latent regression:

$$y_{it}^{*} = \mathbf{X}_{it}^{T}\beta+\varepsilon_{it}; \varepsilon_{it} = u_i+\nu_{it}, u_i \sim \left(0, \sigma_u^2\right), \nu_{it} \sim \left(0, \sigma_\nu^2\right), t = 1, 2, ..., T_i, i = 1, 2, ..., N,$$

$$y_{it} = 1 \text{ if } y_{it}^{*} > 0; 0 \text{ otherwise.}$$

Where, u_i is the unobserved individual specific heterogeneity term. Assume that the person-specific random effect is the same in every period and the unique effects ν_{it} is uncorrelated and independent across periods. The proportion of total variance contributed by the panel level variance component is: $\rho = \dfrac{\sigma_u^2}{\sigma_u^2 + \sigma_\nu^2}$. We set $\sigma_\nu^2 = 1$ for model identification and use Gauss-Hermite quadrature approximation to the log-likelihood to estimate random effects logit model. Intitutively, if r is zero, then the panel level variance component is unimportant and the panel estimator is no different from the pooled estimator. The likelihood ratio test is used to compare the pooled logit estimator with the panel logit estimator. It is found that r is significantly different from zero across all specifications and panel estimator is superior to pooled estimator. The random effects logit estimates for independent living, cohabiting and nursing home and the corresponding odds ratios are presented in Tables 3, 4 and 5, respectively.

The home ownership variable is removed from the institutionalization equation as reported in Table 5, because of the possible problem discussed earlier. In the random effects logit model, the effect of age on both independent living and cohabiting is negatively significant whereas it has a strong positive effect on nursing homes in both specifications. This suggests that age is one of the strong predictors of nursing home entry. Gender shows a clear pattern now. Females are more likely to be institutionalized and less likely to cohabit. Gender is insignificant in the independent living equation.

Loss of a spouse is now positive and significant for both cohabiting and nursing home entry, with the coefficient being larger for the cohabiting equation in both specifications. This suggests that a non-medical event, such as loss of a spouse is a predictor for nursing home entry. Moreover, loss of a spouse decreases independent living, the odds decreases by about 46 to 57 percent.

The effect of education on independent living is positive and that of cohabiting and nursing home is negative. However, as before the magnitudes of education for independent living and institutionalization are reduced after additional variables in the model are added. The odds of educated people being institutionalized or cohabiting is relatively lower than remaining independent. The multinomial logit estimates are corroborated even after individual heterogeneity is controlled for.

The relationship between income and all three forms of living arrangement is found to be non-linear. The relationship between monthly income reported and nursing home is concave and that of independent living and cohabiting is convex. An inverted-U-shaped relationship (i.e., concavity) between income and an institutional form of living arrangement implies that, as income rises to a point, demand for nursing home rises relative to other living arrangements. However, as income rises further, the probability of choosing nursing home falls. This may be suggestive of possible income related inequity for an institutional form of living arrangement. This could indeed be a possibility because there is a minimum daily charge on nursing home use regardless of the level of income.

As before, the effect of home care utilization on nursing home use is negative and on independent living is positive even after controlling for individual heterogeneity. However, the effect of home care utilization on cohabiting turned out to be statistically insignificant, it was marginally significant in the multinomial logit model. This reinforces the result that home care utilization reduces the demand for nursing home and allows the elderly to live independently in the community longer. Home ownership is a significant predictor of both independent living and cohabiting. However, the effect of homeownership on independent living is relatively greater than that of cohabiting.

Similarly, those who lived longer in the community, had better self-reported health status and self-reported general life satisfaction, are more likely to remain in the community and choose independent forms of living arrangement and are less likely to cohabit or be institutionalized.

A number of policy implications can be derived directly from this study. For instance, looking at the random effects logit estimates and ignoring the square term, it can be concluded that as income increases there will be a reduction in the probability of choosing independent living (by about 10%) and cohabiting (in the range of 11 - 13%) whereas there will be an increase in the probability of institutionalization (in the range of 41 to 53%).[8] On the

[8] Interpretation of the non-linear income coefficients can be made by examining the effects at mean income level or some other threshold level of interest.

other hand, provision of formal home care reduces the rate of institutionalization by about 91%. Thus, if the government decides to choose between retirement home subsidy (the effect of which increases disposable income) and formal home care provisions to reduce institutionalization and promote independent living, provision of home care is clearly a superior policy option.

4.3 Fixed Effects Logit Estimates

In this section, we briefly discuss Chamberlin's conditional fixed effects estimates. Under the assumption of conditional independence across cross-section and time periods, the estimates are consistent but inefficient. One virtue of the conditional fixed effects is that unobserved individual specific effects are allowed to correlate with explanatory variables. However, it is to be noted that due to high stability in living arrangements, a large number of observations are dropped out. So, the results need to be interpreted with great caution because the estimates are based on a very restricted data set; based only on 604 individuals for independent living, 345 for cohabiting and 438 for institutional living. Also, in order to estimate fixed effects models, time invariant covariates are dropped out. The results are reported in Table 6.

As in the random effects logit model, the effect of age on both independent living and cohabiting is negative and statistically significant whereas it has a strong positive effect on institutional living arrangement. This again suggests that age is one of the strong predictors of nursing home entry. Like the random effects logit model, loss of a spouse is positive and significant for both cohabiting and institutional living. Further, loss of a spouse is negatively associated with independent living. The relationship between income and all three forms of living arrangement is found to be weak in the fixed effects logit, it is no longer non-linear and insignificant for independent living and cohabiting. Income is positively significant for nursing home entry. This does not preclude possible income related inequity for institutional form of living arrangement.

As before, the effect of home care utilization on nursing home use negative and on independent living is positive even after controlling for unobserved individual heterogeneity. The effect of home care utilization on cohabiting is still negative but statistically insignificant. This again reinforces the previous result that home care reduces the demand for nursing home care. As in the random effects logit model, home ownership is a significant predictor of both independent living and cohabiting. Similarly, those who reported better life satisfaction, are more likely to remain in the community and choose an independent form of living arrangement and less likely to be institutionalized. The health status and years lived in the community are no longer statistically significant.

4.4 Transitions in Living Arangement Decisions

Although the previous sections discuss the determinants of living arrangements, one interesting aspect of longitudinal data is to analyze the transitions in living arrangement decisions. High mortality between the interview years and relatively stable living arrangements for the survivors would not yield a very large sample to analyze all possible shifts. Therefore, we briefly discuss two scenarios here, the shift from an independent living arrangement towards cohabiting and nursing home. Random effects logit estimates are presented in Table 7.

Age has been found to be positively significant for shifting to cohabitation and nursing homes. Loss of a spouse is one of the strongest predictors of the shift to cohabiting. Income has been found to be negative and statistically significant for shifting to nursing home and insignificant for shifting to cohabitation. This again demonstrates that there is an income related constraint on nursing home entry, the odds being decreased by about 13 percent. The effect of home care on shifting from independent living to nursing home is positive and statistically significant. This suggests that those who enter these institutions were home care recipients (i.e., prior home care clients). Alternatively, that the home care service was not sufficient to sustain independent living and eventually led to institutionalization. Home ownership is positively associated with shifts to cohabiting and negatively with nursing home. Similarly, those who report better health status are less likely to shift to cohabiting or nursing home.

5 Conclusions

It is well known that the physical and mental health status of older persons not only depend on the genetic endowments but also on their past lifestyle. Therefore, preventive measures focusing on a healthy lifestyle could delay or deter disease and disability, and allow seniors to live independently in the community. Utilizing Aging In Manitoba (AIM) longitudinal study on 1971 cohort's interview data linking to home care admission data for the AIM study participants, this study for the first time, analyzes the determinants of living arrangement decisions of the elderly Manitobans.

Previous studies on the impact of home care utilization were unclear in the literature and contradictory findings were found in Saskatchewan and British Columbia studies. Our study suggests that provision of formal home care would reduce the demand for nursing home care and enable the elderly to live in the community independently. However, given that home care recipients are more likely to shift to a nursing home, further work is required to draw conclusions on the efficacy of home care program and the degree of substitutability between home care and nursing home care.

After controlling for individual heterogeneity, it is found that loss of a spouse affects independent living negatively and both cohabiting and nursing home positively. The effect of age on nursing home residence is positive and on independent living and cohabiting is negative. Females are more likely to be institutionalized than males. Educated people are more likely to live independently than to cohabit or be institutionalized. Similarly, those who are healthy and satisfied in life are more likely to live independently instead of cohabiting or entering nursing homes. Those who lived longer in the community are more likely to live independently or cohabit rather than enter an institution. Home ownership is positively associated with both independent living and cohabiting. There is some evidence of an income gradient to institutionalization. In the random effects logit model, a concave relationship between monthly income and nursing homes is found whereas for independent living and cohabiting the relationship is convex.

There are some limitations of this study that can be improved upon in future research works. Our results are valid subject to survival. Using the mortality data from Manitoba Health, duration dependence models would provide better insight into the predictors of survival and mortality. In this study, we used the 1971 cohort of AIM study sample. It would be interesting to use a more recent panel, say the 1983 cohort, to see if we obtain similar results to corroborate our findings.

References

[1] Amemiya, T. (1981), "Qualitative response models: a survey", *Journal of Economic Literature*, 19(4), 481-536.

[2] Amemiya, T. (1985), "Advanced Econometrics", Cambridge, Harvard University Press.

[3] Arrow, K. J. (1963), "Uncertainty and the welfare economics of medical care", *American Economic Review*, 53, 941-973.

[4] Ben-Akiva, M., and Lerman, S. R. (1985), "Discrete Choice Analysis: Theory and Application to Travel Demand", Cambridge, MA, The MIT Press.

[5] Black, C., Roos, N.P., Havens, B. and MacWilliam, L. (1994), "Rising use of physician services by the elderly: the contribution of morbidity", *Canadian Journal on Aging*, 14(2), 225-244.

[6] Börsch-Supan, A. H. (1990), "A dynamic analysis of household dissolution and living arrangement transition by elderly Americans", in David A. Wise ed., Issues in the Economics of Aging, 89-114, The University of Chicago Press.

[7] Börsch-Supan, A. H., Hajivassiliou, V., Kotlikoff, L. J., and Morris, J. N. (1992), "Health, children, and elderly living arrangements: a multiperiod multinomial probit model with unobserved heterogeneity and autocorrelated errors", in David A. Wise, ed., Topics in the Economics of Aging, 79-107, The University of Chicago Press.

[8] Carriére, Y., and Pelletier, L. (1995), "factors underlying the institutionalization of elderly persons in Canada", *Journal of Gerontology:Social Sciences*, 50B(3), S164-S172.

[9] CIHI (2000), "Health care in Canada: a first annual report", Canadian Institute for Health Information, Ottawa,Ontario.

[10] CIHI (2001), "Development of national indicators and reports for home care: final project report, Canadian Institute for Health Information, Ottawa, Ontario.

[11] Cramer, J. S., and Ridder, G. (1991), "Pooling states in the multinomial mlogit model", *Journal of Econometrics*, 47, 267-272.

[12] Davidson, J., and Mackinon, J. (1993), "Estimation and Inference in Econometrics", New York, Oxford University Press.

[13] Di Matteo, L. and Di Matteo, R. (2001), "Public homecare expenditures in Canada", *Canadian Public Policy*, 27(3), 313-333.

[14] Denton, F. T., and Spencer, B. G. (1983), "Population aging and future health costs in Canada", *Canadian Public Policy*, 9(2), 155-163.

[15] Denton, F. T., and Spencer, B. G. (1995), "Demographic change and the cost of publicly funded health care", *Canadian Journal on Aging*, 14(2), 174-192.

[16] Denton, F. T., and Spencer, B. G. (1997), "Population aging and the maintenance of social support systems", *Canadian Journal on Aging*, 16(3), 485-498.

[17] Denton, F. T., and Spencer, B. G. (1999), "Social and economic dimensions of an aging population", SEDAP Research Papers, McMaster University, Hamilton, Canada.

[18] Engers, M., and Stern, S. (2002), "Long-term care and family bargaining", *International Economic Review*, 43(1), 73-114.

[19] Forbes, W. F., Jackson, J. A., and Kraus, A. S. (1987), "Institutionalization of the Elderly in Canada", Butterworths, Toronto, Ontario.

[20] Freedman, V.A. (1996), "Family structure and the risk of nursing home admission", *Journal of Gerontology B: Psychological Sciences and Social Sciences*, 51(2), S61-S69.

[21] Garber, A. M., and Phelps, C. E. (1997), "Economic foundations of cost-effectiveness analysis", *Journal of Health Economics*, 16, 1-31.

[22] Greene, W. H. (1997), "Econometric Analysis", 3rd edition, Prentice Hall.

[23] Grundy, E. and Glaser, K. (1997), "Trends in, and transitions to, institutional residence among older people in England and Wales, 1971-91", *Journal of Epidemiology and Community Health*, 51(5), 531-540.

[24] Grundy, E., 1999. Changing role of the family and community in providing support for the elderly. In: Cliquet, R., Nizamuddin, M (Eds.), Population Aging: Challenges for Policies and Programmes in Developed and Developing Countries. United Nations Population Fund/Population and Family Study Centre, New York, pp. 103-122.

[25] Hausman, J., and McFadden, D. L. (1984), "Specification tests for the multinomial logit model", *Econometrica*, 52(5), 1219-1240.

[26] Hausman, J., Leonard, G. K., and McFadden, D. L. (1995), "A utility consistent, combined discrete choice and count data model assessing recreational use losses due to natural resource damage", *Journal of Public Economics*, 56,1-30.

[27] Health Canada (1997), "Toward healthy-aging communities: a population health approach", a report by Linda Macleod and associates, Health Canada, Ottawa, Ontario.

[28] Health Canada (1999), "Provincial and territorial home care programs: a synthesis for Canada", Health Canada, Ottawa, Ontario.

[29] Health Canada (1999), "Report on the national roundtable on home and community care", Health Canada, Ottawa, Ontario.

[30] Health Services Utilization and Research Commission (2000), The Impact of Preventive Home Care and Seniors Housing on Health Outcomes", Saskatoon, SK, Canada.

[31] Heckman, J. (1979), "Sample selection as a specification error", *Econometrica*, 47, 153-62.

[32] Hoerger, T. J., Picone, G. A., and Sloan, F. A. (1996), "Public subsidies, private provision of care and living arrangements of the elderly", *The Review of Economics and Statistics*, 78(3), 428-440.

[33] Hollander, M. (2001), "Final Report of the Study on the Comparative Cost Analysis of Home Care and Residential Care Services. Vancouver, BC: Hollander Analytical Services Ltd.

[34] Houle, C., Berthelot, J., David, P., Mustard, A. C., Roos, L., and Wolfson, M. C. (1996), "Project on matching Census 1986 database and Manitoba health care files: private household component", Statistics Canada, Ottawa, Ontario.

[35] Hum, D., and Simpson, W. (1996), "Canadians with disabilities and the labour market", *Canadian Public Policy*, 22, 285-289.

[36] Hum, D., and Simpson, W. (2002), "Disability onset among aging Canadians: evidence from panel data", *Canadian Journal on Aging*, 21(1), 117-136.

[37] Jones, A. M. (2000), "Health econometrics", in J. P. Newhouse and A. J. Culyer, eds., Handbook of Health Economics, volume 1a, 265-364, Elsevier.

[38] Kane, R.L., Chen, Q., Finch, M., Blewett, L., Burns, R., and Moskowitz, M. (2000), "The optimal outcomes of post-hospital care under medicare", *Health Services Research*, 35(3), 615-661.

[39] Kenkel, D. S. (1995), "Should you eat breakfast? Estimates from health production functions", Health Economics, 4, 15-29.

[40] Kotlikoff, L. J. and Morris, J. N. (1990), "Why don't the elderly live with their children? A new look", in David A. Wise., ed., Issues in the Economics of Aging, 149-169, The University of Chicago Press.

[41] Lavery, L.A., Van Houtum, W. H., and Armstrong, D.G. (1997), "Institutionalization following diabetes-related lower extremity amputation", *American Journal of Medicine*, 103(5), 383-388.

[42] Management Committee of Cabinet (1977), Program Encyclopedia: Manitoba Health Services Commission, Winnipeg, MB, Canada.

[43] Manitoba Health (2001), "Manitoba Health Annual Repot", 1999/2000. Manitoba Health. Winnipeg. 2001.

[44] McFadden, D. L. (1974), "Conditional logit analysis of qualitative choice behavior", in Frontiers in Econometrics edited by Paul Zarembka, New York, Academic Press.

[45] McFadden, D. L. (1981), "Econometric models of probabilistic choice", in Structural Analysis of Discrete Data. Edited by C. F. Manski and D. McFadden, Cambridge, MIT Press.

[46] McFadden, D. L. (1984), "Qualitative response models in Z. Griliches and M. Intriligator", Handbook of Econometrics, Amsterdam, North Holland.

[47] McFadden, D. L. (1989), "A method of simulated moments for estimation of discrete response models without numerical integration", *Econometrica*, 57(5), 995-1026.

[48] McFadden, D. L. (2001), "Economic choices", The American Economic Review, 91(3), 351-378.

[49] McFadden, D. L., and Ruud, P. A. (1994), "Estimation by Simulation", *The Review of Economics and Statistics*, 76(4), 591-608.

[50] Menec, V. H., MacWilliam, L., Soodeen, R., and Mitchell, L. (2002), "The health and health care use of Manitoba's seniors: have they changed over time? ", Manitoba Centre for Health Policy, Department of Community Health Sciences, University of Manitoba, Winnipeg, MB, Canada.

[51] Mustard, A. C., Finlayson, M., Derksen, S., and Berthelot, J. (1999), "What determines the need for nursing home admission in a universally insured population?", *Journal of Health Services Research & Pol icy*, 4(4), 197-203.

[52] Norton E.C. (2000), "Long-term care", in A. J. Culyer and J. P. Newhouse, eds., Handbook of Health Economics, volume 1, Elsevier, 956-994.

[53] Rockwood, K., Stolee, P., and McDowell, I. (1996), "Factors associated with institutionalization of older people in Canada: testing a multifactorial definition of frailty", *Journal of the American Geriatric Society*, 44(5), 578-582.

[54] Roos, N.P., Shapiro, E., Bond, R., Black, C., Finlayson ,G., and Newburn-Cook, C. (2001a), "Changes in Health and Health Care Use of Manitobans: 1985- 1998", Manitoba Centre for Health Policy, Department of Community Health Sciences, University of Manitoba, Winnipeg, MB, Canada.

[55] Roos, N. P., Stranc, L., Peterson, S., Mitchell, L., Bogdanovic, B., and Shapiro, E. (2001b), "A look at home care in Manitoba", Manitoba

Centre for Health Policy, Department of Community Health Sciences, University of Manitoba, Winnipeg, Canada.

[56] Roos, N. P., Mitchell, L, Peterson, S., and Shapiro, E. (2001c), "Perspectives on home care data requirements", Manitoba Centre for Health Policy, Department of Community Health Sciences, University of Manitoba, Winnipeg.

[57] Rosenberg, M. (2000), "The effects of population aging on the Canadian health care system", SEDAP Research Papers, McMaster University, Hamilton, Ontario.

[58] Smith, G.E., Koimen, E., and O'Brien, P.C. (2000), "Risk factors for nursing home placement in a population-based dementia cohort", *Journal of the American Geriatric Society*, 48(5), 519-525.

[59] Steinbach, U. (1992), "Social networks, institutionalization, and mortality among elderly people in the United States", *Journal of Gerontology*, 47(4), S183- S190.

[60] Stern, S. (1995), "Estimating family long-term care decision in the presence of endogenous child characteristics", *Journal of Human Resources*, 30(3), 551-580.

[61] Tinker, A., Askham, J., Hancock, R., Mueller, G., Stuchbury, R. 85 Not Out: A study of people aged 85 and over at home. Anchor Trust, Oxford, 2001.

[62] Tomiak, M., Berthelot, J., and Mustard, A. C. (1998), "A profile of the health care utilization of the disabled population in Manitoba, *Medical Care*, 36(9), 1383-1397.

[63] Tomiak, M., Berthelot, J., and Mustard, A. C. (2000), "Factors associated with nursing home entry for elders in Manitoba, Canada", *Journal of Gerontlogy, A: Biological Sciences and Medical Sciences* 55A(5), M279-M287.

[64] Trottier, H., Martel, L., Houle, C., Berthelot, J., and Légaré, J. (2000), "Living at home or in an institution: what makes the difference for seniors?", *Health Report*, 11(4), 49-61.

Appendix A
Table A.1
Living Arrangements of the Elderly

Living Arrangement*	1971	Survivors in 1983	Survivors in 1990	Survivors in 1996
Independent	3,145 (65.48%)	1,042 (68.64%)	423 (67.14%)	104 (48.6%)
Cohabiting	1,049 (21.84%)	156 (10.28%)	44 (6.98%)	13 (6.07%)
Nursing home	609 (12.68%)	320 (21.08%)	163 (25.87%)	97 (45.33%)
Total	4,803 (100%)	1,518 (100%)	630 (100%)	214 (100%)

* Independent living arrangement is defined as living alone or living with spouse; cohabiting is defined as living with children, siblings, friends, parents, or grand children.

Table A.2
Age Distribution at Baseline 1971

Age	<= 65	66-70	71-75	76-80	81-85	86-90	91-95	96-100	>100
Count	186	1,418	1,101	886	682	385	111	24	10
Percent	3.87	29.52	22.92	18.45	14.20	8.02	2.31	0.50	0.21

Table A.3
Educational Status at Baseline 1971

Education (in years)	0	1-4	5-8	9-10	11-12	13-16	>16	Missing
Count	401	705	1797	694	497	161	49	499
Percent	8.35	14.68	37.41	14.45	10.35	3.35	1.02	10.39

Table A.4
Marital Status

Marital Status	1971	Survivors in 1983	Survivors in 1990	Survivors in 1996
Single	508 (10.6%)	148 (9.75%)	68 (10.79%)	26 (12.15%)
Married	2,417 (50.3%)	552 (36.36%)	153 (24.29%)	34 (15.89%)
Widowed*	1,874 (39.0%)	818 (53.89%)	407 (64.60%)	154 (78.96%)
Total	4,803 (100%)	1,518 (100%)	630 (100%)	214 (100%)

* Includes divorced and separated cases. The relatively small frequencies do not permit us to report them in separate categories as per the ethics guidelines.

Note: The total observations include missing cases.

Table A.5
Health Status

Self Reported Health Status	**1971**	**Survivors in 1983**	**Survivors in 1990**	**Survivors in 1996**
Excellent	563 (11.7%)	119 (7.84%)	43 (6.83%)	9 (4.21%)
Good	2,073 (43.2%)	597 (39.33%)	182 (28.89%)	58 (27.10%)
Fair	1,266 (26.4%)	389 (25.63%)	133 (21.11%)	27 (12.62%)
Poor/ Bad	439 (9.17%)	126 (8.30%)	45 (6.33%)	12 (3.74%)
Not Applicable	462 (9.62%)	287 (18.91%)	227 (36.03%)	112 (52.34%)
Total	4,803 (100%)	1,518 (100%)	630 (100%)	214 (100%)

Table A.6
General Life Satisfaction

Self Reported Health Status	**1971**	**Survivors in 1983**	**Survivors in 1990**	**Survivors in 1996**
Excellent	751 (15.64%)	223 (14.69%)	64 (10.16%)	10 (4.67%)
Good	2,617 (54.49%)	637 (41.96%)	256 (40.63%)	74 (34.58%)
Fair	735 (15.30%)	235 (15.48%)	70 (11.11%)	22 (10.28%)
Poor/Bad	134 (2.79%)	47 (3.09%)	9 (1.43%)	-
Not Applicable	566 (11.78%)	376 (24.77%)	231 (36.67%)	108 (50.47%)
Total	4,803 (100%)	1,518 (100%)	630 (100%)	214 (100%)

Table A.7
Years lived in the Community

Number of Years	**1971**	**Survivors in 1983**	**Survivors in 1990**	**Survivors in 1996**
> 50/all life	2,557 (53.24%)	605 (39.86%)	246 (39.05%)	87 (40.65%)
26-50		367 (24.18%)	143 (22.70%)	52 (24.30%)
11-25	1,149 (23.92%)	276 (18.18%)	89 (14.13%)	31 (14.49%)
6-10	387 (8.06)	117 (7.71%)	44 (6.98%)	16 (7.48%)
3-5	325 (6.77%)	73 (4.81%)	37 (5.87%)	9 (4.21%)
0-2	361 (7.52)	57 (3.75%)	40 (6.35%)	15 (7.01%)
Total	4,803 (100%)	1,518 (100%)	630 (100%)	214 (100%)

Note: The total number of observation includes missing cases.

Table A.8 Average Monthly Reported Income

Reported Income from all Sources*	**1971**	**Survivors in 1983**	**Survivors in 1990**	**Survivors in 1996**
Average	187.1008	500.76	376.2	547.93
Minimum	0	0	0	0
Maximum	1,701	4,254	6,900	3,500
Standard Deviation	135.943	325.6	524.2	542
Total Observations	4,767	1,486	594	210

* The sources of income are from: a) private pensions, pension from private companies, wages, salary income from business, farm, professional practice, rents, interests, dividends and insurance annuities; b) Old Age Security (OAS), Guaranteed Income Supplements (GIS), War Veterans Allowance/Pension, Social Allowance, Public welfare Agency, Unemployment Insurance, Canada Pension Plan, Old Age Assistance, Manitoba Supplement for Pensioners, Tax credits; and c) financial assistance from children, relatives, friends, Church, Service groups, private agency, etc.

Table A.9
Home Ownership

Own house	**1971**	**Survivors in 1983**	**Survivors in 1990**	**Survivors in 1996**
Yes	2,637 (54.9%)	709 (46.71%)	206 (32.7%)	49 (22.9%)
No/NA	2,166 (45.1%)	809 (53.29%)	424 (67.3%)	165 (77.1%)
Total	4,803 (100%)	1,518 (100%)	630 (100%)	214 (100%)

Table 1
Descriptive Statistics (Panel 1971-96)

Variable	1971 Mean	1971 Std. Dev.	1983 Mean	1983 Std. Dev.	1990 Mean	1990 Std. Dev.	1996 Mean	1996 Std. Dev.
Independent living	.657	.475	.693	.4613	.683	.465	.495	.501
Cohabiting	.217	.413	.104	.3058	.069	.254	.062	.241
Nursing home	.126	.331	.202	.4020	.247	.432	.443	.498
Female	.529	.499	.593	.4915	.663	.473	.714	.453
Home care received	0	0	.095	.294	.515	.50	.505	.501
Age in years	75.2	7.46	82.6	4.833	88.05	3.781	92.9	2.82
Widows/separated/ divorced	.389	.488	.536	.4988	.64	.481	.719	.45
Home ownership	.552	.497	.472	.4994	.337	.473	.236	.425
Lived >25 years	.535	.499	.649	.4773	.651	.477	.662	.474
Lived 11-25 years	.241	.427	.185	.3885	.145	.3522	.147	.356
Lived 6-10 years	.081	.273	.078	.2684	.074	.262	.076	.266
Lived 3-5 years	.068	.252	.049	.2162	.062	.242	.043	.203
Lived 0-2 years	.075	.264	.038	.1921	.067	.251	.071	.258
Excellent health	.118	.322	.080	.2715	.072	.259	.043	.203
Good health	.434	.496	.396	.4893	.298	.458	.271	.446
Fair health	.264	.441	.259	.4383	.22	.415	.129	.335
Poor health	.083	.276	.071	.2575	.054	.226	.029	.167
Bad health/ not reported	.101	.301	.193	.3949	.355	.479	.529	.50
Excellent life	.157	.364	.148	.3553	.104	.306	.048	.213
Good life	.547	.498	.423	.4943	.424	.495	.352	.479
Fair life	.154	.361	.156	.3631	.113	.317	.105	.307
Poor life	.023	.149	.024	.1538	.012	.108	0	0
Bad life/ not reported	.119	.324	.247	.4318	.347	.476	.49	.501
Education: 0-4 yrs	.231	.422	.242	.4286	.18	.385	.19	.394
Education: 5-10 yrs	.521	.499	.5	.5002	.508	.500	.5	.501
Education: >=11 yrs	.148	.355	.195	.3964	.149	.357	.243	.429
Income/100	1.87	1.36	5.01	3.26	3.76	5.24	5.48	5.42
Number of Observations	4767		1486		594		210	

Note: We do not have the same individuals in all periods due to deaths from one survey to the next.

Table 2

Multinomial Logit Estimates of Living Arrangement (Pooled from 1971-96)

	Model 1		Model 2	
	(1)	(2)	(3)	(4)
	Cohabiting	Nursing home	Cohabiting	Nursing home
Age in years	0.001	0.110*	0.001	0.082*
	(0.005)	(0.006)	(0.005)	(0.007)
Female	-0.119***	0.214**	-0.115	-0.021
	(0.071)	(0.084)	(0.072)	(0.099)
Married/single (ref.)				
Widows/separated/divorced	0.617*	0.540*	0.648*	0.091
	(0.074)	(0.086)	(0.077)	(0.099)
Less than 5 years of or no education (ref.)				
Education: 5-10 years	-0.343*	-1.291*	-0.254*	-0.886*
	(0.072)	(0.081)	(0.075)	(0.098)
Education: >=11 years	-0.473*	-1.162*	-0.365*	-0.651*
	(0.104)	(0.113)	(0.108)	(0.137)
Income/100	-0.079*	0.234*	-0.057*	0.254**
	(0.022)	(0.078)	(0.022)	(0.099)
Income square/10000	0.001*	-0.013***	0.001**	-0.013
	(0.000)	(0.008)	(0.000)	(0.010)
year83	-0.668*	-0.605*	-0.751*	-0.548*
	(0.111)	(0.113)	(0.115)	(0.135)
year90	-1.298*	-0.737*	-1.134*	-0.120
	(0.183)	(0.151)	(0.201)	(0.206)
year96	-1.001*	-0.302	-0.848**	0.338
	(0.310)	(0.225)	(0.339)	(0.309)
Home ownership			0.022	-6.355*
			(0.076)	(0.707)
Home care received			-0.461**	-1.630*
			(0.207)	(0.199)
Years lived the community: 10 years or less (ref.)				
Lived >25 years			0.347*	-0.431*
			(0.097)	(0.103)
Lived 11-25 years			0.245**	-0.678*
			(0.109)	(0.128)
Self reported health status: Fair/Poor/bad/not applicable (ref.)				
Excellent health			-0.253**	-0.677*
			(0.122)	(0.225)
Good health			-0.116	-0.568*
			(0.074)	(0.106)
Self reported general life satisfaction: Fair/Poor/bad/not applicable (ref.)				
Excellent life			-0.280**	-1.697*
			(0.113)	(0.171)
Good life			-0.244*	-1.392*
			(0.080)	(0.102)
Constant	-0.990**	-10.227*	-1.063**	-5.771*
	(0.401)	(0.489)	(0.422)	(0.561)
Observations	7,046	7,046	7,044	7,044
Log likelihood	-5440.13		-4525.36	
Cragg & Uhler's R2	0.221		0.447	

Robust standard errors in parentheses

*** significant at 10%; ** significant at 5%; * significant at 1%

Table 3

Determinants of Independent Living: Random Effects Logit Estimates

	(1) Coefficient	(2) Odds Ratio	(3) Coefficient	(4) Odds Ratio
Age in years	-0.059*	0.942*	-0.036*	0.965*
	(0.006)		(0.006)	
Female	-0.040	0.961	0.056	1.058
	(0.084)		(0.087)	
Married/single (ref.)				
Widows/separated/divorced	-0.844*	0.430*	-0.618*	0.539*
	(0.088)		(0.090)	
Less than 5 years of or no education (ref.)				
Education: 5-10 years	1.069*	2.911*	0.674*	1.963*
	(0.087)		(0.090)	
Education: >=11 years	1.118*	3.058*	0.652*	1.920*
	(0.120)		(0.124)	
Income/100	-0.088*	0.916*	-0.090*	0.914*
	(0.026)		(0.027)	
Income square/10000	0.005*	1.005*	0.004*	1.004*
	(0.002)		(0.002)	
year83	0.898*	2.456*	0.938*	2.555*
	(0.110)		(0.119)	
year90	1.035*	2.815*	0.633*	1.883*
	(0.149)		(0.177)	
year96	0.102	1.107	-0.220	0.802
	(0.222)		(0.257)	
Home ownership			1.163*	3.200*
			(0.091)	
Home care received			1.438*	4.213*
			(0.184)	
Years lived the community: 10 years or less (ref.)				
Lived >25 years			0.119	1.126
			(0.096)	
Lived 11-25 years			0.233**	1.263**
			(0.113)	
Self reported health status: Fair/Poor/bad/not applicable (ref.)				
Excellent health			0.409*	1.505*
			(0.144)	
Good health			0.317*	1.373*
			(0.086)	
Self reported general life satisfaction: Fair/Poor/bad/not applicable (ref.)				
Excellent life			1.113*	3.043*
			(0.135)	
Good life			0.939*	2.557*
			(0.091)	
Constant	5.210*		1.962*	
	(0.462)		(0.480)	
Log likelihood	-4110.46		-3823.2	
ρ	0.391 (Chi2=162.61)		0.384 (Chi2=136.59)	
	(0.009)		(0.01)	
Observations	7,046 (id=4,767)		7,044 (id=4,767)	

Standard errors in parentheses

*** significant at 10%; ** significant at 5%; * significant at 1%

Table 4
Determinants of Cohabiting - Random Effects Logit Estimates

	(1) Coefficient	(2) Odds Ratio	(3) Coefficient	(4) Odds Ratio
Age in years	-0.035*	0.966*	-0.024*	0.977*
	(0.008)		(0.008)	
Female	-0.265**	0.767**	-0.211***	0.810***
	(0.119)		(0.120)	
Married/single (ref.)				
Widows/separated/divorced	0.923*	2.518*	1.096*	2.993*
	(0.125)		(0.132)	
Less than 5 years of or no education (ref.)				
Education: 5-10 years	-0.117	0.889	-0.155	0.856
	(0.115)		(0.121)	
Education: >=11 years	-0.273***	0.761***	-0.314***	0.731***
	(0.166)		(0.173)	
Income/100	-0.133*	0.875*	-0.112*	0.894*
	(0.031)		(0.031)	
Income square/10000	0.002**	1.002**	0.002**	1.002**
	(0.001)		(0.001)	
year83	-0.925*	0.397*	-1.076*	0.341*
	(0.166)		(0.173)	
year90	-1.614*	0.199*	-1.627*	0.196*
	(0.253)		(0.288)	
year96	-1.665*	0.189*	-1.701*	0.183*
	(0.418)		(0.446)	
Home ownership			0.614*	1.847*
			(0.120)	
Home care received			-0.233	0.792
			(0.275)	
Years lived the community: 10 years or less (ref.)				
Lived >25 years			0.720*	2.054*
			(0.144)	
Lived 11-25 years			0.571*	1.770*
			(0.162)	
Self reported health status: Fair/Poor/bad/not applicable (ref.)				
Excellent health			-0.250	0.779
			(0.188)	
Good health			-0.049	0.952
			(0.117)	
Self reported general life satisfaction: Fair/Poor/bad/not applicable (ref.)				
Excellent life			0.034	1.034
			(0.176)	
Good life			0.057	1.058
			(0.123)	
Constant	0.515		-1.290***	
	(0.612)		(0.665)	
Log likelihood	-3037.39		-2999.98	
ρ	0.588 (Cho2=216.04)		0.588 (Chi2=209.06)	
	(0.01)		(0.011)	
Observations	7,046 (id=4,767)		7,044 (id=4,767)	

Standard errors in parentheses
*** significant at 10%; ** significant at 5%; * significant at 1%

Table 5

Determinants of Institutionalization: Random Effects Logit Estimates

	(1) Coefficient	(2) Odds Ratio	(3) Coefficient	(4) Odds Ratio
Age in years	0.227*	1.255*	0.180*	1.197*
	(0.024)		(0.018)	
Female	0.475*	1.608*	0.328**	1.389**
	(0.181)		(0.158)	
Married/single (ref.)				
Widows/separated/divorced	0.947*	2.579*	0.692*	1.997*
	(0.190)		(0.165)	
Less than 5 years of or no education (ref.)				
Education: 5-10 years	-2.541*	0.079*	-1.289*	0.276*
	(0.261)		(0.185)	
Education: >=11 years	-2.392*	0.091*	-0.894*	0.409*
	(0.313)		(0.235)	
Income/100	0.348*	1.417*	0.429*	1.535*
	(0.048)		(0.054)	
Income square/10000	-0.014*	0.986*	-0.018*	0.982*
	(0.003)		(0.004)	
year83	-0.611*	0.543*	-0.806*	0.447*
	(0.196)		(0.190)	
year90	-0.351	0.704	0.185	1.203
	(0.246)		(0.268)	
year96	1.071*	2.918*	1.213*	3.365*
	(0.379)		(0.399)	
Home care received			-2.419*	0.089*
			(0.312)	
Years lived the community: 10 years or less (ref.)				
Lived >25 years			-1.980*	0.138*
			(0.206)	
Lived 11-25 years			-1.931*	0.145*
			(0.239)	
Self reported health status: Fair/Poor/bad/not applicable (ref.)				
Excellent health			-1.176*	0.309*
			(0.315)	
Good health			-0.895*	0.409*
			(0.174)	
Self reported general life satisfaction: Fair/Poor/bad/not applicable (ref.)				
Excellent life			-2.676*	0.069*
			(0.335)	
Good life			-2.177*	0.113*
			(0.211)	
Constant	-21.785*		-14.997*	
	(2.165)		(1.461)	
Log likelihood	-2436.47		-2076.64	
ρ	0.765 (Chi2=182.70)		0.656 (Chi2=108.11)	
	(0.013)		(0.016)	
Observations	7,046 (id=4,767)		7,046 (id=4,767)	

Standard errors in parentheses

*** significant at 10%; ** significant at 5%; * significant at 1%

Table 6

Determinants of Living Arrangements (1971-96): Fixed Effects Logit Estimates

	(1) Independent	(2) Cohabiting	(3) Nursing Home
Age in years	-0.043*	-0.106*	1.270*
	(0.014)	(0.021)	(0.044)
Married/single (ref.)			
Widows/separated/divorced	-0.919*	1.312*	1.569***
	(0.199)	(0.282)	(0.822)
Income/100	0.014	0.023	0.261**
	(0.022)	(0.031)	(0.122)
Home ownership	1.069*	0.809*	
	(0.186)	(0.271)	
Home care received	1.288*	-0.436	-1.799**
	(0.221)	(0.310)	(0.798)
Years lived the community: 10 years or less (ref.)			
Lived >25 years	0.162	0.107	-0.546
	(0.179)	(0.296)	(0.678)
Lived 11-25 years	0.159	0.033	-1.234
	(0.211)	(0.332)	(0.825)
Self reported health status: Fair/Poor/bad/not applicable (ref.)			
Excellent health	0.101	0.102	-0.301
	(0.266)	(0.384)	(1.368)
Good health	0.247	-0.221	-0.859
	(0.156)	(0.237)	(0.729)
Self reported general life satisfaction: Fair/Poor/bad/not applicable (ref.)			
Excellent life	1.229*	0.081	-3.176**
	(0.238)	(0.339)	(1.497)
Good life	0.801*	0.195	-0.948
	(0.154)	(0.239)	(0.613)
year83	1.015*	-0.458***	-12.256*
	(0.161)	(0.236)	(0.599)
year90	0.857*	-0.060	-17.773
	(0.223)	(0.351)	(0.000)
Log likelihood	-429.38	-208.64	-33.66
Observations	1,616 (id=604)	905 (id=345)	1,171 (id=438)

Standard errors in parentheses
*** significant at 10%; ** significant at 5%; * significant at 1%

Table 7
Determinants of Transitions in Living Arrangements (1971-96)
(Random Effects Logit Estimates)

	(1) Independent to Cohabiting	(2) Odds Ratio	(3) Independent to Nursing Home	(4) Odds Ratio
Age in years	0.357**	1.429**	0.204*	1.226*
	(0.146)	(2.44)	(0.052)	(3.91)
Married/single (ref.)				
Widows/separated/divorced	3.893*	49.050*	0.224	1.251
	(1.432)	(2.72)	(0.193)	(1.16)
Income/100	0.130	1.139	-0.136*	0.873*
	(0.151)	(0.86)	(0.044)	(3.08)
Home care received	-0.796	0.451	0.924*	2.520*
	(1.306)	(0.61)	(0.338)	(2.73)
Home ownership	3.424*	30.700*	-0.827*	0.437*
	(1.290)	(2.66)	(0.264)	(3.14)
Self reported health status: Fair/Poor/bad/not applicable (ref.)				
Excellent health	-4.190**	0.015**	-0.728**	0.483**
	(1.782)	(2.35)	(0.333)	(2.19)
Good health	-0.581	0.559	-0.201	0.818
	(0.872)	(0.67)	(0.201)	(1.00)
Self reported general life satisfaction: Fair/Poor/bad/not applicable (ref.)				
Excellent life	0.463	1.589	-0.427	0.653
	(1.140)	(0.41)	(0.283)	(1.51)
Good life	-0.716	0.489	-0.625*	0.535*
	(0.915)	(0.78)	(0.241)	(2.60)
year83	-3.092**	0.045**	-1.959*	0.141*
	(1.492)	(2.07)	(0.416)	(4.71)
year90	-7.001**	0.001**	-1.947*	0.143*
	(3.010)	(2.33)	(0.500)	(3.89)
Constant	-42.726*		-15.073*	
	(14.459)		(3.854)	
Log likelihood	-265.46		-726.64	
ρ	0.959 (Chi2=6.03)		0.356 (Chi2=2.16)	
	(0.007)		(0.08)	
Observations	1,377 (id=966)		1,652 (id=1131)	

Standard errors in parentheses
*** significant at 10%; ** significant at 5%; * significant at 1%

A: Transition from Independent to Cohabiting	1971-83	1983-90	1990-96	Total	B: Transition from Independent to Institution	1971-83	1983-90	1990-96	Total
0	833	377	94	1,304	0	833	377	94	1,304
1	52	18	3	73	1	199	90	59	348
Total	885	395	97	1,377	Total	1,032	467	153	1,652

0: Refers to the total number of survivors of 1971 cohort remained in an independent living arrangement.
1: Panel A refers to the number of survivors of 1971 cohort who shifted from an independent living arrangement to cohabiting and panel B refers to those who shifted from an independent living arrangement to institutions.

Conclusions

It is well known that understanding the underlying process of the demand for health and health care utilization is crucial for a better assessment of the role of public intervention in the health sector. This issue is gaining momentum in both developed and developing countries alike. In particular, governments, public policy makers, economists, and citizens around the world are debating who should pay for what and how best to organize and deliver health services so as to allocate scarce resources efficiently and work towards a healthier society. In addition, the secular rise in spending on health care, relative to other goods and services, coupled with onset of aging in the population raise important issues in designing health policy. This thesis has explored several theoretical and empirical aspects of these issues using unexplored and unique data sets, appropriate economic theoretical frameworks, and advanced methodological tools in each of the three essays.

Essay 1 examines the factors determining the utilization of different types of health care from recent Canadian National Population Health Survey conducted by Statistics Canada. It uses the number of visits to GPs, specialists, and dentists and the number of nights spent in hospital as measures of utilization of health care. An intuitively appealing economic framework, in which individuals maximize the net benefits of visits, is used to base the analysis of health care utilization. Several techniques, namely Negative Binomial Models, Hurdle Models, Zero Inflated Models and Latent Class Models are used in this essay to analyze health care utilization. However, the latent class modelling framework suggests that it is a superior statistical technique if the data permits modelling unobserved heterogeneity and overdispersion.

This essay addresses two fundamental issues about the analysis of health care utilization in a publicly funded health care system. First, what is an appropriate framework to analyze health care utilization using data on individual citizens that is now commonly available? Secondly, can this framework be used to address crucial policy issues? Can analysis of health care utilization provide an indication of whether there is supplier-induced demand? How health care utilization responds to need? Do income and supplemental health insurance matter in health care utilization?

It is found that the decision to contact a health professional (i.e., *ex ante* utilization) and the decision about how much to utilize, proxied by the number of visits (i.e., *ex post* utilization), are essentially two distinct stochastic processes requiring two-stage models of utilization. The most striking results from this essay is that supplemental health insurance increases outpatient health visits, that there is vertical equity in the utilization of health care, and that there are some indications of supplier induced demand for health care.

Future research can improve our understanding of the factors underlying health care utilization by improving the data and techniques used. In terms of the data, important variables missing in the survey data include waiting time, travel time, and out of pocket spending for each visit to different types of health professionals. The absence of these variables limits our ability to identify some of the interesting and crucial parameters relating to the demand for health care. Future research would be a technical improvement in dealing with the endogenous regressors. One plausible endogeneity problem could be that the self reported health status indicators may themselves be determined by the other regressors in the model, such as life style variables.

Essay 2 examines the issues relating to demand for outpatient health care in rural India. In India, much of the health services are typically provided at little or no *monetary* cost. Although there exists extensive public support for hospitals, medical education and drugs in India, it is not so clear is whether governments spent appropriately in order to raise access to and use of health care regardless of ability to pay. This essay addresses the following set of interrelated questions. What are the determinants of demand? How important are price, income, quality, and access in health care provider choice? How do rich and poor individuals make decisions about their treatment in response to price?

This essay employs a discrete choice model to explain the underlying determinants of the demand for outpatient health care in rural India based on National Sample Survey (NSS) data for the first time. As opposed to fixed choice sets used in the literature, a variable choice set is constructed and used in this study to reflect the true choice generating process as close as possible. The relevant price data for unchosen alternatives in the choice set are imputed. The paper discusses econometric methods relating to identification, scaling, invariance, and consistency with the utility maximization hypothesis that underlies the basis of modelling health care demand.

Contrary to many earlier studies on the demand for health care in developing countries, prices and income are found to be statistically significant determinants of health care choice. Distance is a pronounced inhibiting factor in the demand for outpatient health care in rural India. Another set of inhibiting factors governing health seeking behaviour are those which are attributed to stylized facts of the socioeconomic environment - especially those who adopted bad habits and those who do not have access to safe drinking water and latrine facilities. The result is suggestive of gender bias in the demand for health care.

Although this essay extended to a variable choice set and imputed price data, one weakness of this study is that the analysis was restricted by the available data quality. In order to effectively use a discrete choice model of health care demand, it is important that data on quality and other characteristics of alternatives need to be collected along with the survey data. Although we tried with different imputed data sets and obtained similar results, the estimated results would greatly improve with precision in the presence of a high quality data. Nevertheless, the present study is a preliminary attempt in the direction of modelling demand for health care in a typical developing country and would motivate future research in this area.

Aging of the population - coupled with the onset of disability, widowhood, and deterioration in health status among the elderly - may lead to revised decisions about the living arrangements of the elderly, which in turn has a bearing on demand for long-term care. Although elderly Canadians are generally healthy, there is uncertainty about demand for certain types of services, including health, medical, and personal care. Uncertainty arises not only due to incidence of illness, disability, and widowhood but also because the provision of different types of care is indeterminate. Linking Aging In Manitoba (AIM) longitudinal study on 1971 cohort's interview data with home care admission data, essay 3 explores the underlying determinants of elderly living arrangements. This essay uses pooled multinomial logit and random effects logit model to analyze the determinants of elderly living arrangement decisions.

It is found that home care admission (*ex ante* home care utilization) reduces the demand for nursing home and increases the demand for independent living. Loss of a spouse affects independent living negatively and both cohabiting and nursing home residence positively. The effect of age on nursing home residence is positive and on independent living and cohabiting is negative. Educated people are more likely to live independently than cohabit or enter an institution. Similarly, those who are healthy and satisfied in life are more likely to live independently instead of cohabiting or entering nursing homes. Those who lived longer in the community are more likely to live independently or cohabit rather than enter an institution. Home ownership is positively associated with both independent living and cohabiting. The results are suggestive of possible income related inequity in institutionalization.

However, there are a number of ways of improving this initial work on elderly living arrangement decisions. Although the living arrangement decisions of the elderly are relatively stable, a fraction of people do change their initial living arrangements over the period of retired lifetime. So, it would be interesting to incorporate these inherent dynamic futures into the living arrangement decisions. Basically, one needs to take into account the possible transitions of living arrangements in the life course of the elderly within a multinomial logit setting. The second aspect is that most of the elderly died over these long period, causing censoring from one survey to another. Further, moving into a nursing home is conditioned by a rationing rule. The expert panel screens all potential applicants and if the composite health condition of an applicant exhibits a threshold limit then the person is allowed to enter a nursing home residence. A discrete-time hazard rate model to account for sample attrition and living arrangement transitions within a multinomial logit setting is an agenda for future research.