ENDOGENOUS SWITCHING OF INSURANCE REGIME AND THE DEMAND FOR HEALTH CARE SERVICES AMONG THE INSURED AND UNINSURED

Şevket Alper KOÇ[1]
Çağatay KOÇ[2]

ABSTRACT

Whether one has health insurance not only determines one's out-of-pocket health care costs, but also one's access to health care. This paper estimates the demand for a *portfolio* of health care services by insurance regime to examine the determinants of health care demand among the insured and uninsured. Such an analysis would not be possible using traditional estimation methods given the endogeneity of insurance, the change of insurance regime and the discreteness and the non-negativity of the dependent variable. These aspects motivate the application of an endogenous switching model for count data, also known as a type-5 Tobit, using the 1993 National Health Interview Survey. Certain significant relationships hold across insurance status. For example, employed individuals demand fewer operations and nights in the hospital. Certain significant relationships only appear in one insurance state. For example, among the insured, the higher educated individuals demand more specialist visits. In the demand for physician services, health status is more influential among the uninsured than the insured. This relationship is reversed in the demand for hospital services, where health is less influential among the uninsured than the insured. These results suggest that insurance likely modifies the relationship between the independent variables and health care utilization by providing an access to an entirely different system of care.

ÖZET

Bir kişinin sağlık sigortasının olması sadece sağlık hizmetlerine harcadığı parayı değil aynı zamanda sağlık hizmetlerine ulaşabilmesini de etkiler. Bu çalışma sigortalılar ile sigortalı olmayanlar arasındaki sağlık hizmetleri talebinin belirleyicilerini incelemek için sağlık hizmetleri talebini sigorta rejimleri ile tahmin etmektedir. Sigortanın içselliği ve bağımlı değişkenin pozitifliği ve kesikli değişken olmasından dolayı bu tür bir analiz geleneksel tahmin yöntemleri ile yapılamamaktadır. Bundan dolayı bu çalışmada 1993 Milli Sağlık Araştırma Anketi kullanılarak type-5 Tobit modeli olarak da bilinen içsel switching model kullanılmıştır. Sağlık statüleri arasında belli anlamlı ilişkiler görülmüştür. Örneğin, çalışan bireyler daha az ameliyat ve hastanede daha az gece yatma talebinde bulunmaktadırlar. Bazı anlamlı ilişkiler ise sadece bir sigorta türünde görünmektedir. Örneğin, sigortalılar arasında daha yüksek eğitimli bireylerin daha fazla uzman vizitesi talep ettikleri gözlemlenmiştir. Doktor hizmetleri talebinde ise sağlık statüsünün sigortalı olmayanlar arasında sigortalılara göre daha etkili olduğu gözlemlenmiştir. Hastane hizmetleri talebi açısından ise bu ilişkinin ters olduğu gözlemlenmiştir. Bu sonuçlar göstermektedir ki sigorta bağımsız değişkenler ile sağlık hizmetlerinden faydalanma arasındaki ilişkiye etkileyebilmektedir.

[1] Asst. Professor, Kocaeli University, Department of Economics, Umuttepe, Izmit, Turkey, alperkoc@kocaeli.edu.tr.
[2] Asst. Professor, Department of Economics, University of Texas at Arlington, Arlington, TX 76019, ckoc@uta.edu

1. Introduction

In the United States, health insurance has dual role. It provides membership to the insured's health care system and reduces out-of-pocket health care costs through third party payment. While much of the insurance literature focuses on comparing health care utilization between health care systems among the insured, we study the gap between the insured and uninsured. Using hypothetical scenarios, Mort *et al*. (1996) show that physicians' prescribing patterns differ based on the presence of insurance suggesting that even when the insured and uninsured receive care from the same providers, the uninsured are set apart into an alternative health care regimen. In this paper, we estimate the demand for a *portfolio* of health care services by insurance regime to examine the determinants of health care demand among the insured and uninsured. (Mort, Edwards, Emmons, Convery and Blumenthal 1996: 783).

In estimating the demand for health care services, we control for the possible endogeneity in the choice of health insurance. Most studies of health care utilization recognize the endogeneity of health insurance; however, the literature has generally focused on comparing utilization between types of insurance (e.g. fee-for-service versus health maintenance organization) or level of out-of-pocket payment (e.g. RAND Health Insurance Experiment).[3] Dowd *et al.* (1991) incorporate the endogenous switching of health care system in their estimation of health care utilization. They compare health care utilization among employees insured in fee-for-service system and health maintenance organizations in the Minneapolis – St. Paul area.[4] Certainly, differences between managed care and fee-for-service health care systems exist, as found by Dowd *et al*. (1991), but these differences are narrow compared to the chasm separating the uninsured and insured. In this paper, we compare the demand for a *portfolio* of health care services among the insured and uninsured by estimating a count version of endogenous switching model using a nationally representative dataset. (Dowd, Feldman, Cassou and Finch 1991: 85).

Our analysis would not be possible using traditional econometric methods given the endogeneity of insurance, the change of insurance regime and the discreteness and non-negativity of the dependent variable. These aspects motivate the application of an endogenous switching model for count data. The econometric methods applied for this analysis are founded on Terza's (1998) general formulation of a count data regression framework that accommodates both endogenous treatment effects and endogenous regime switching. The endogenous treatment effect version of his general formulation has been used in various applied studies to analyze the effect of endogenous vehicle ownership on trip frequency, (Terza 1998: 129) the impact of chronic illicit drug use on emergency room utilization (McGeary and French 2000: 153), and the effect of physician advice on alcohol consumption (Kenkel and Terza 2001: 165). We expand on the range of applications of Terza's (1998) general formulation by providing an example to the endogenous regime switching, also known as a type-5 Tobit model, version of his methodology.

---

[3] In the path breaking RAND Health Insurance Experiment study, participants were randomly assigned health insurance policies varying by copayment and deductible (Newhouse *et al.* 1993). All were insured to a maximum out-of-pocket expense of $1,000 and most were covered under the same type of health care system.

[4] As part of the RAND Health Insurance Experiment, Manning *et al.* (1984) randomly assigned participants to a health maintenance organization or fee-for-service plan. However, they modeled the change in regime using an intercept shift instead of a switching model. (Manning, Leibowitz, Goldberg, Rogers and Newhouse 1984: 1505)

We proceed as follows: Section 2 presents the data and explains the formulation of the health care service counts. This section also contains the working definition of health. Section 3 delineates the econometric methodology. Section 4 presents results dealing with health insurance demand, endogeneity of insurance, and the demand for health care services among the insured and uninsured. Section 5 concludes.

2. The Data

We study a sample of 18,116 adults between the ages of 18 and 64 drawn from the 1993 National Health Interview Survey and its Health Insurance Supplement compiled by the U.S. Department of Health and Human Services. Since the insurance supplement is available only for the 3rd and 4th quarters, we start with a data set of 37,225 individual records. 11,412 observations containing veterans, individuals who served in Armed Forces and their families are removed since their health care utilization and access to health care distinctly differs from the general population. 7,684 additional observations are removed from the data set due to incomplete responses and another 13 are removed due to inconsistent responses. Each observation contains 33 variables relevant to the estimation of health insurance and health care service demand. The definitions and summary statistics of these variables are reported in Tables 1 and 2. Fourteen of these variables relate to the personal background of each individual- such as age, sex, race, education, family size, etc.- and affect health care and health insurance demand through preferences and risk factors.

Seven variables relate to the individual's ability to purchase health insurance; these include income, employment and access to group insurance variables.[5] Seven variables relate to the individual's health status. Following Cameron *et al*. (1988) and Vera-Hernandez (1999), we apply a practical definition of health based on definitions of disease, illness and disability. (Cameron, Trivedi, Milne and Piggott 1988: 85, Vera-Hernandez 1999: 579). For each of these three dimensions of health, we construct a criterion for a binary indicator. *Disease*, the state when something has objectively and demonstratively gone wrong in the mechanics of an organism, is identified by the presence of chronic health conditions. *Illness*, the mental state that is known to follow physical and psychological disorders, is inherently subjective. The usage of “Poor” and “Fair” responses in the self-reported health index as the indicator of illness allows for this subjectivity. *Disability*, the final component of health, is the loss of opportunity imposed by social and environmental factors on people with impairments relative to a perceived norm of hindrance. We identify the presence of disability based on activity limitation status. According to these definitions, disability is a subset of disease since impairment is necessary for disability and disease is necessary for impairment. The combinations of these three indicators separate the sample into six disjoint health groups, H1-H6. The creation of these groups removes concerns about multicollinearity between disease, illness, and disability variables. In addition to these six health groups, LIMDY, the total restricted activity days in the past two weeks, is included as a health variable to account within group variation caused by acute conditions. Such conditions

[5] For the income variable, we transform discrete interval responses to a single income variable. We avoid inherent biases created by arbitrarily choosing points within the intervals by applying an expected value estimation technique that utilizes the log-normal distribution assumption of income. The details of this technique are available upon request.

affect an individual's short-run health care demand, and their anticipation may also influence the expected value of health insurance.

Table 1. Definitions of Explanatory and Health Insurance Variables

Variable	Description of Variable	N=18,116	
		Mean	St. Dev.
sex	1 if male	0.45	0.49
age	Number of years old	36.93	11.31
age2	Age squared divided by 1,000	1.49	0.89
white	1 if white	0.80	0.39
imm	1 if born outside of the United States	0.13	0.33
mar	1 if married	0.64	0.47
nhs	1 if no high school diploma	0.16	0.36
univ	1 if high school graduate and some years beyond high school	0.47	0.49
rgn1	1 if resides in the Northwest	0.19	0.39
rgn2	1 if resides in the Midwest	0.26	0.44
rgn3	1 if resides in the South	0.31	0.46
rgn4	1 if resides in the West	0.22	0.41
famsz	Number of family members	3.009	1.49
city	1 if Metropolitan Statistical Area (MSA) in excess of 100,000	0.75	0.42
inc	Estimated family income divided by 1,000 measured in 1993 dollars	46.67	39.32
pubins	1 if publicly insured	0.08	0.27
privins	1 if privately insured	0.73	0.44
empl1	1 if employed in the public sector	0.11	0.31
empl2	1 if employed in the private sector	0.64	0.47
empl	1 if employed	0.75	0.43
offer	1 if a family member is offered private insurance by an employer	0.88	0.31
H1	1 if disease, illness and disable	0.05	0.23
H2	1 if disease, disable and no illness	0.07	0.26
H3	1 if disease, illness and no disable	0.02	0.14
H4	1 if illness, no disease and no disable	0.02	0.14
H5	1 if disease, no illness and no disable	0.20	0.40
H6	1 if no disease, no illness and no disable	0.61	0.48
limdy	Number of restricted activity days over the last two weeks	0.59	2.33

As for the dependent variables, we study five counts of health care services. Two are counts of hospital services over the preceding year: operations (OPER) and nights in the hospital (HOSP). The OPER count includes only non-pregnancy-related operations conducted in a hospital, and not therapeutic or diagnostic non-operative procedures. HOSP contains only visits to non-military, short-stay hospitals since individuals with access to military facilities are removed. The last three dependent variables are counts of physician services over the preceding two weeks: visits with a physician (MDMS), visits with a specialist (SPMS), and visits with a general practitioner (GPMS). MDMS includes only in-person consultations with a physician and is the sum of SPMS and GPMS. These consultations may be conducted in a home, clinic or office, but not in a laboratory.[6]

Table 2. Definitions of Utilization Variables

Variable	Description of Variable	Mean	St. Dev.	Percentage of Zeros
mdms	Number of visits with a physician over the last two weeks	0.126	0.44	90.08
hosp	Number of nights in a short-stay hospital over the last year	0.394	2.38	92.48
oper	Number of operations over the last year	0.061	0.27	94.70
gpms	Number of visits with a general practitioner over the last two weeks	0.0657	0.27	95.07
spms	Number of visits with a specialist over the last two weeks	0.068	0.33	94.64

3. Econometric Analysis

Cameron *et al*. (1988) provide a model of individual behavior relevant to decisions about health care and insurance, which provides a setting for the estimation of the interdependent demands of health care and insurance. We rely on Cameron *et al*.'s (1988) economic model and focus on the estimation of the demand for health care services among the insured and uninsured.

Three attributes complicate the estimation of the interdependent demands for insurance and health care utilization. First, health insurance is not distributed randomly. It is likely endogenous to health care utilization leading to potential biases in the estimation of health care utilization if left uncontrolled. Second, the differences in health care utilization across insurance regimes cannot be addressed with a single parameter. Insurance likely modifies the relationship between the socioeconomic variables and health care utilization by providing access to an entirely different system of care. Third, utilization of health care is discrete and non-negative in the form of a count of services over a period of time. These three attributes lead us to apply the

[6] GPMS may be interpreted as the utilization of primary care, while SPMS is the utilization of health care services that involve specialized clinical expertise.

endogenous switching model (a.k.a. type-5 Tobit) proposed by Terza (1998). The remainder of this section motivates the application of the type-5 Tobit.

3.1. Endogeneity Problem

Unobservable factors that influence one's utilization of health care might be correlated with the unobservables that determine insurance demand. If such endogeneity is left uncontrolled in the estimation of health care utilization, the resulting parameter estimates will be biased.

Agents who enter into a health insurance contract are rarely selected at random. Agent characteristics, such as health status, may influence the decision to enter into a contract and thus create a self-selection bias. If these characteristics can be hidden pre-contractually, the resulting contract may adversely affect the uninformed parties in the contract. This phenomenon is known as *adverse selection*. In this case, the unobservables in the utilization regression and unobservables that determine the outcome of the binary private insurance variable are positively correlated. In addition, insurance companies may attempt to prohibit high-risk individuals from entering their risk pool, a procedure known as *screening*. In this case, the unobservables are negatively correlated. Either selection bias, adverse selection or screening, potentially confounds the estimation of health care service demand.

Recognizing these potential biases, we apply estimation techniques that control for the potentially non-random distribution of insurance. However, these techniques require the estimation of the relationship between insurance demand and its determinants. Given the binary nature of private insurance status, we apply a simple probit model. As for the determinants of insurance, we include variables that also determine the utilization of health care services, such as health, employment, education and income; and variables that are independent of the use such as whether a family member is offered insurance by their employer, immigration status, number of family members and the region in which the individual resides. Therefore, we are able to identify the endogenous variable using determinants that are independent of health care use and the non-linearity in functional form between the probit and the count model of health care use. These two aspects aid in controlling for the endogeneity of insurance by deterring the potential reverse causality.

3.2. Selectivity into Regimes

Ignoring the change in regime, one might suggest an endogenous treatment model (i.e., a type-3 Tobit model) for count data[7], or apply a generalized method of moments (GMM) estimation with the predicted probability of having private insurance from the probit model as an instrument for insurance.[8] Both procedures would control for the potential bias caused by the endogeneity of insurance and each allows the estimation of the determinants of health care utilization. However,

[7] For applications of this model in various contexts using Terza's (1998) general methodology, see Terza (1998), McGeary and French (2000), and Kenkel and Terza (2001). (Terza 1998:129, McGeary and French 2000: 153, Kenkel and Terza (2001: 165)

[8] Examples of GMM estimation of an endogenous treatment effect model appear in Vera-Hernandez (1999) and Schellhorn (2001). Vera-Hernandez (1999) analyzes the implications of duplicate coverage on demand for visits to specialist, and Schellhorn (2001) studies the effect of the choice of deductibles in the mandatory basic health insurance in Switzerland on physician service utilization. (Vera-Hernandez 1999: 579, Schellhorn 2001: 441)

neither allows for the complete interaction between insurance and the determinants of utilization. Therefore, our approach is to apply the endogenous switching model (i.e., a type-5 Tobit model) for count data proposed by Terza (1998).[9] With this approach, we are able to estimate the demand equations for the insured and the uninsured separately and identify differences in utilization across insurance regimes. The remainder of this section provides a brief description of the type-5 Tobit model. The model's formal details appear in Terza (1998).

Let $f\left(y_j^*\middle|X_j,\varepsilon_j\right)$, $j=1,2$, be the conditional probability density function of the count dependent variables. $y_j^*=0,1,2,...$ are the count dependent variables, which represent the number of health care services over a stated period of time for sub-population j. For the specific application in this paper, y_1^* is the utilization of health care by the insured and y_2^* is the utilization of health care by the uninsured. ε_j is the interpersonal heterogeneity[10] component, and X_j is a vector of explanatory variables.

Switching variable d is characterized by a latent variable model $d^*=Z\alpha+\xi$ where Z is a vector of exogenous variables determining d^*. We assume that only the sign of d^* is observed and thus define

$$d=\begin{cases}1 & \text{if } Z\alpha+\xi>0\\ 0 & \text{otherwise.}\end{cases}$$

d is the outcome of the binary switching variable, which represents the presence of private health insurance.

Instead of fully specifying the model by assuming a particular form for $f\left(y_j^*\middle|X_j,\varepsilon_j\right)$, it is assumed that the conditional means can be written as

$$E\left[y_j^*\middle|X_j,\varepsilon_j\right]=\exp\left\{X_j'\beta_j+\varepsilon_j\right\},\ \text{ for } j=1,2\,,$$

which is a much weaker assumption. Note that this conditional mean assumption is satisfied in both the Poisson and Negative Binomial regression models, which are the benchmark models for count data. The following sample selection rule is followed:

$$y_1=\begin{cases}y_1^* & \text{if } d^*>0,\\ 0 & \text{otherwise,}\end{cases}\quad \text{and}\quad y_2=\begin{cases}y_2^* & \text{if } d^*\le 0,\\ 0 & \text{otherwise.}\end{cases}$$

[9] As stated in the introduction, Terza's (1998) original formulation accommodates both endogenous switching and endogenous treatment effect models. (Terza 1998:129)

[10] ε_j reflects a specification error, such as unobserved omitted exogenous variables from the set X_j.

Finally, we assume that $\{\varepsilon_1, \varepsilon_2, \xi\}$ are *i.i.d.* drawings from a trivariate normal distribution with mean vector zero and covariance matrix

$$\begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{11}\rho_1 \\ \sigma_{12} & \sigma_{22}^2 & \sigma_{22}\rho_2 \\ \sigma_{11}\rho_1 & \sigma_{22}\rho_2 & 1 \end{bmatrix}.$$

To find the regression functions for the sub-populations, one needs to derive

(1) $$E[y_1 \mid X_1, d=1] \text{ and } E[y_2 \mid X_2, d=0].$$

In this paper's application, these two equations specify the expected utilization of health care conditional on being insured and uninsured, respectively. Terza (1998) shows that

(2) $$E\left[y_1 \middle| X_1, d=1\right] = \exp\left\{X_1'\beta_1^*\right\}\frac{\Phi\left(\theta_1 + Z'\alpha\right)}{\Phi\left(Z'\alpha\right)},$$

where β_1^* is the same as β_1 except that the first element corresponding to the constant term is multiplied by $\frac{\sigma_{11}}{2}$, $\theta_1 = \sigma_{11}\rho_1$, and $\Phi(.)$ denotes the standard normal cumulative distribution function. Similarly

(3) $$E\left[y_2 \middle| X_2, d=0\right] = \exp\left\{X_2'\beta_2^*\right\}\frac{1-\Phi\left(\theta_2 + Z'\alpha\right)}{1-\Phi\left(Z'\alpha\right)},$$

where $\theta_2 = \sigma_{22}\rho_2$, and β_2^* is defined similarly as above.

Given the conditional mean functions in (2) and (3), one can motivate the exponential regression functions,

$$y_j = m_j\left(X_j, Z, \beta_j^*, \theta_j; \alpha\right) + e_j \text{ for } j = 1, 2,$$

where $m_j\left(X_j, Z, \beta_j^*, \theta_j; \alpha\right)$ is defined in (2) for $j=1$ and it is defined in (3) for $j=2$, and e_j is a stochastic error with $E(e_j \mid X_j, d) = 0$. To estimate the above system, we employ the two-stage method of moments estimation proposed in Terza (1998).

In the first stage of this type-5 Tobit model, we estimate the determinants of insurance demand employing a simple probit analysis as discussed. In the second stage, we split the sample into two by insurance status and estimate the expected utilization of health care services among the insured and uninsured separately;[11] the complete interaction in the demand equations (Lee 1978:

[11] As stated in Terza (1998), a heteroskedasticity-consistent estimator is used for the asymptotic covariance matrix of the second-stage estimator. (Terza 1998:129)

415). Within the functional forms, this procedure corrects for the endogeneity of insurance by multiplying the expected demand functions with terms similar to the inverse Mill's ratio. This procedure is a non-linear least squares analog to the Heckman estimation procedure (Heckman 1979: 153). The sign and significance of the theta (θ_j) estimate corresponds to the correlation between the unobservables in the demand for insurance and those in the demand for health care and is indicative of the endogeneity bias.

For an independent variable that is common in the vector of explanatory variables in stages 1 and 2 equations, its marginal effect includes both the direct effect on the dependent variable and the indirect effect through the term similar to the inverse Mill's ratio. We derive the marginal effect estimators for the two stage method of moments estimation taking into account both of these effects. The sign of the marginal effect depends both on the sign and magnitude of β, and those of α.[12]

4. Results

4.1. Demand for Health Insurance

To estimate the endogenous switching model, we estimate a simple probit model on the presence of private health insurance, which also allows us to evaluate the determinants of insurance demand. The results, which appear in Table 3, suggest that demographic characteristics have relatively small marginal effects on the demand for health insurance as compared to variables representing health, employment, and access to group coverage.

Employment increases the demand for insurance by decreasing premiums. Federal tax incentives motivate employers to subsidize insurance premiums of their employees (Gruber 2000: 646). The public sector has historically provided more generous employee benefits, which may explain the difference between the employment coefficients. The offering of health insurance through place of employment or that of a family member, the OFFER variable, signifies access to group insurance.[13] This indicator of access has the largest marginal effect, increasing the probability of private insurance by over 34 percent. Access to group insurance decreases insurance premiums due to the vast economies of scale in the insurance industry. Last, higher income significantly increases demand for insurance. Individuals with high income may purchase health insurance to protect their assets from health-induced expenditure shocks. These results are consistent with those of earlier studies. Cameron *et al*. (1988) find that income is an important determinant of health insurance decision and Vera-Hernandez (1999) finds that income and employment status are very important in determining health insurance choice.

Under the hypothesis of health-related adverse selection, individuals who anticipate poor health are more likely to purchase private insurance due to the corresponding increase in the expectation of health care utilization. The probit results clearly cast doubt on the existence of adverse selection for observable variables. For example, being in the worst health state (H1) decreases the predicted probability of private insurance by 16 percent compared to being in the state of perfect

---

[12] Derivations are available upon request.

[13] It is important to note that the access to group insurance has no direct effect on the demand for health care services; the OFFER variable is not included in the demand estimations.

health (H6). All the estimated health coefficients are significantly negative except those of H3, H5 and LIMDY. The coefficient estimates of H3 and LIMDY are negative, but insignificant, and that of H5 is positive, but also insignificant. These estimated coefficients suggest that the structure of the insurance industry favors the healthy likely through screening procedures.

Table 3. Probit Regression for Private Insurance

N=18,116	Coefficient	t-stat	Marginal Effects
sex	0.005	0.19	0.001
age	-0.034	-4.92*	-0.008
age2	0.636	7.18*	0.164
white	0.224	7.42*	0.061
mar	0.355	12.42*	0.096
famsz	-0.192	-19.99*	-0.049
imm	-0.245	-6.72*	-0.068
city	-0.036	-1.26	-0.009
rgn2	0.069	1.83	0.017
rgn3	-0.196	-5.60*	-0.052
rgn4	-0.214	-5.68*	-0.058
nhs	-0.373	-11.02*	-0.107
univ	0.285	10.15*	0.073
H1	-0.526	-9.27*	-0.163
H2	-0.185	-4.06*	-0.051
H3	-0.100	-1.24	-0.027
H4	-0.281	-3.78*	-0.081
H5	0.035	1.10	0.009
limdy	-0.003	-0.66	-0.0009
emp1	0.742	13.90*	0.142
emp2	0.310	10.32*	0.083
offer	1.020	25.55*	0.341
inc	0.014	30.29*	0.003
constant	-0.331	-2.59*	N.A.[a]
L	-6975.879[b]		

Note: We estimate a logit regression and find similar results. All estimated coefficients of the logit model have the same sign and only small differences in their t-statistics. * indicates statistical significance at 5 percent.
[a] Not Applicable.
[b] ***L*** is the log likelihood of the Probit regression. The percent of observations correctly predicted is 83.29.

4.2. Endogeneity and Screening

In addition to these insurance demand results, we also obtain an estimate for the correlation between the unobservable determinants of insurance and health care demands, *theta*, as part of the estimation of health care service demands. As Coulson *et al*. (1995) and Terza (1998) point out, this correlation not only allows one to test the endogeneity of insurance, it also provides evidence as to the type of selection. According to the results in Table 4, *theta* is significantly negative in the demand for specialist visits and physician visits among the insured. These results are consistent with *screening* in the demand for specialist and physician visits. All other *theta* estimates are insignificantly negative, except one; the correlation in the uninsured's demand for operations is insignificantly positive.

Table 4. Endogeneity Tests by Category

N=18,116	MDMS	GPMS	SPMS	HOSP	OPER
t-stat of Theta: Insured	-2.87*	-1.04	-2.29*	-0.01	-0.85
t-stat of Theta: Uninsured	-0.85	-0.61	-0.5	-0.28	1.37

Note: *** indicates significance at 5 percent.**

This evidence of screening in the self selection results may reflect the structure of the U.S. health care industry. Most privately insured individuals obtain their insurance coverage through employment or that of a family member; employment and insurance are bundled goods. Employees who choose the same health plan usually pay the same premium and are subject to the same other requirements. Therefore, insurance companies do not charge premiums based on individually rated expected cost. Instead, insurers have a financial incentive to control the use of costly services, such as specialty services, by adjusting their risk pool (i.e., screening). In addition, under screening, the endogeneity of insurance would appear most prominently among services with high direct costs to the insurer.

4.3. Health Care Utilization among the Insured and Uninsured

The parameter estimates produced by the switching model illustrate the determinants of health care utilization (See Tables 5-9). The non-health related coefficients portray a complex relationship between the demand for health care services and the insurance state.

Certain significant relationships, described in this paragraph, hold across insurance states. Males demand less physician services, but more nights in the hospital than females. White individuals demand fewer general practitioner visits than non-whites. In the demand for operations, the coefficient of AGE is negative and the coefficient of AGE2 is positive. Initially, as people get older their utilization of operations decrease, but by age 50, the utilization increases with age suggesting a u-shaped relationship. Married individuals demand more operations. Employed individuals demand fewer operations and nights in the hospital. In fact, being employed decreases the demand for operations by 56 percent. For all health care services and insurance states, the

coefficient of income is insignificant, suggesting that income does not influence the demand for health care services. It is interesting to compare this result with some previous findings. Cameron *et al*. (1988) find that health care service use appears to be income-insensitive. Schellhorn (2001) reports that income does not have a significant effect on both specialist and primary physician visits, (Schellhorn 2001: 441) whereas Vera-Hernandez (1999) finds that income significantly increases the demand for specialist visits.
Certain significant relationships only appear in one insurance state. Among the insured, residing in the city increases the demand for physician services, the higher educated individuals demand more specialist visits, employed individuals demand more general practitioner visits and being white increases the demand for nights in the hospital. It is important to note that these results only hold for the insured. Schellhorn (2001) finds that residing in the city increases the demand for specialist visits and both Vera-Hernandez (1999) and Schellhorn (2001) find that the higher educated individuals demand more specialist visits. However, since they model the demand for health care services with an endogenous treatment model (i.e., modeling the change in insurance regime with an intercept shift), the effect of an independent variable on health care use is the same for both the insured and uninsured. Our results suggest that insurance modifies the relationship between the independent variables and health care utilization.

Among the uninsured, males demand fewer operations and public insurance has a large and positive effect on the demand for all services.[14] For example, public insurance increases the demand for operations by 73 percent among the uninsured.

Health as a determinant of health care service demand has the greatest marginal effect on all services, except operations. In the demand for the three physician services, all health coefficients, except one, are significantly positive, independent of the insurance state. The coefficient for H4, which represents individuals who are ill but have no disease or disability, is insignificant. Thus, just being ill does not cause one to go to the doctor. In the demand for hospital care, all health coefficients are significantly positive, independent of insurance state. In the demand for operations, all health coefficients, except for that of H4, are significantly positive among the insured. These same coefficients among the uninsured, however, are insignificant and negative, except for H5, which is significant and positive. In other words, health does influence the insured's demand for operations, but does not appear to affect the uninsured's demand. The last health variable, LIMDY, the number of days of limited activity, significantly increases the demand for all health care services, as expected.

These results are generally consistent with those of Cameron *et al*. (1988) who report that the health status measures are statistically much more significant in explaining health care utilization than socio-economic characteristics. However, by comparing the marginal effects of the health indicators across insurance status, our results suggest another important difference between the insured and uninsured health care service demand. For the three physician counts, the marginal effects of health are less among the insured than among the uninsured. Thus, in the demand for physician services, health state is more influential among the uninsured than the insured. This

[14] We recognize that some privately insured adults are also covered by public insurance, but we omit this binary variable from the insured regressions, because its variation is too small, with mean less than 1 percent, to allow convergence of our optimization routine.

relationship is reversed in the demand for hospital services, where health is less influential among the uninsured than the insured.

5. Concluding Remarks

Whether one has health insurance not only determines one's out-of-pocket health care costs, but also one's access to health care. In this paper, we estimate the demand for a *portfolio* of health care services by insurance regime to examine the determinants of health care demand among the insured and uninsured using the 1993 National Health Interview Survey and its Health Insurance Supplement. Such estimation would not be possible using traditional methods given the endogeneity of insurance, the change of insurance regime and the discreteness and non-negativity of the dependent variable. These aspects motivate the application of an endogenous switching model for count data founded on Terza's (1998) general formulation of a count data regression that accommodates both endogenous treatment effects and endogenous regime switching.

We find that certain significant relationships hold across insurance status. For example, males demand less physician services, but more nights in the hospital than females. Employed individuals demand fewer operations and nights in the hospital. We also find that income does not have a significant effect for all health care services and insurance status. Certain significant relationships, however, only appear in one insurance state. For example, among the insured, residing in the city increases the demand for physician services and the higher educated individuals demand more specialist visits. These results suggest that insurance likely modifies the relationship between the independent variables and health care utilization by providing access to an entirely different system of care.

We find that the health status measures are statistically much more significant in explaining health care utilization than socio-economic characteristics. However, by comparing the marginal effects of the health indicators across insurance status, our results suggest another important difference between the insured and uninsured health care service demand. In the demand for physician services, health state is more influential among the uninsured than insured. This relationship is reversed in the demand for hospitals services, where health is less influential among the uninsured than the insured.

The econometric technique we employ requires the estimation of the relationship between insurance demand and its determinants. Our results cast doubt on the existence of health-related adverse selection for observable variables. However, we find evidence for the endogeneity of insurance in the demand for specialist visits and physician visits, both among the insured. For these services, our results suggest that unobservable determinants of insurance demand and health care service use are negatively correlated, consistent with the screening hypothesis.

References

Cameron, A. C., P. K. Trivedi, F. Milne and J. Piggott (1988) “A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia,” **Review of Economic Studies**, 55, 85-106.

Coulson, N. E., J. V. Terza, C. A. Neslusan and B. C. Stuart (1995) “Estimating the Moral-Hazard Effect of Supplemental Medical Insurance in the Demand for Prescription Drugs by the Elderly,” **American Economic Review**, 85, 122-126.

Dowd B., R. Feldman, S. Cassou and M. Finch (1991) "Health Plan Choice and the Utilization of Health Care Services," **Review of Economics and Statistics**, 73, 85-93.

Gruber J. (2000) "Health Insurance and the Labor Market," in **Handbook of Health Economics**, Culyer A.J. and J.P. Newhouse (eds.), North-Holland: Amsterdam, 646-706.

Heckman J.J. (1979) "Sample Selection Bias as a Specification Error," **Econometrica**, 47, 153-162.

Kenkel D.S. and J.V. Terza (2001) "The Effect of Physician Advice on Alcohol Consumption: Count Regression with an Endogenous Treatment Effect," **Journal of Applied Econometrics**, 16, 165-184.

Lee L.F. (1978) "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables," **International Economic Review**, 19, 415-433.

Manning W.G., A. Leibowitz, G.A. Goldberg, W.H. Rogers and J.P. Newhouse (1984) "A Controlled Trial of the Effect of a Prepaid Group Practice on Use of Services," **New England Journal of Medicine**, 310, 1505-1510.

McGeary K.A. and M.T. French (2000) "Illicit Drug Use and Emergency Room Utilization," **Health Services Research**, 35, 153-169.

Mort E.A., J.N. Edwards, D.W. Emmons, K. Convery and D. Blumenthal (1996) "Physician Response to Patient Insurance Status in Ambulatory Care Clinical Decision-Making: Implications for Quality of Care," **Medical Care**, 34, 783-797.

National Center for Health Statistics (U.S.) 1993. National Health Interview Survey and Health Insurance Supplement. The Center: Hyattsville, MD.

Newhouse J.P. and the Insurance Experiment Group (1993) **Free for All? Lessons From the RAND Health Insurance Experiment**, Harvard University Press: Cambridge, MA.

Schellhorn M. (2001) "The Effect of Variable Health Insurance Deductibles on the Demand for Physician Visits," **Health Economics**, 10, 441-456.

Terza J.V. (1998) "Estimating Count Data Models with Endogenous Switching: Sample Selection and Endogenous Treatment Effects," **Journal of Econometrics**, 84, 129-154.

Vera-Hernandez A.M. (1999) “Duplicate Coverage and Demand for Health Care: The Case of Catalonia,” **Health Economics**, 8, 579-598.

Table 5. Two-Stage Method of Moments Estimation for Physician Visits

	INSURED			UNINSURED		
Variables	**Coef.**	**t-stats**	**% Eff.**	**Coef.**	**t-stats**	**% Eff.**
sex	-0.366	-3.51*	-0.365	-0.381	-2.63*	-0.380
age	-0.047	-1.56	-0.055	0.026	0.88	0.018
age2	0.548	1.52	0.702	-0.394	-1.13	-0.261
white	-0.123	-0.79	-0.069	-0.018	-0.13	0.028
mar	-0.148	-0.86	-0.062	0.030	0.23	0.104
univ	0.136	1.44	0.205	-0.029	-0.19	0.030
city	0.311	2.63*	0.302	0.068	0.52	0.060
inc	0.000	0.21	0.003	-0.001	-0.17	0.002
empl	0.171	1.17	0.171	-0.035	-0.18	-0.034
H1	1.313	5.52*	1.186	2.116	10.20*	2.006
H2	0.805	4.06*	0.760	1.355	6.61*	1.316
H3	1.129	5.06*	1.104	1.902	7.16*	1.881
H4	0.253	0.65	0.184	0.411	1.09	0.352
H5	0.777	5.49*	0.785	1.293	7.89*	1.300
limdy	0.129	13.13*	0.128	0.055	5.34*	0.054
pubins				0.520	3.48*	0.520
constant	-1.544	-2.15*		-3.942	-5.43*	
Theta	-0.613	-2.87*		-0.382	-0.85	

Note: % Eff. refers to the marginal percentage effects for the corresponding variables. * indicates significance at 5 percent.

Table 6. Two-Stage Method of Moments Estimation for General Practitioner Visits

	INSURED			UNINSURED		
Variables	**Coef.**	**t-stats**	**% Eff.**	**Coef.**	**t-stats**	**% Eff.**
sex	-0.205	-1.91**	-0.205	-0.542	-4.52*	-0.541
age	-0.014	-0.51	-0.017	0.037	0.82	0.030
age2	0.200	0.58	0.259	-0.467	-0.87	-0.332
white	-0.259	-1.87**	-0.238	-0.316	-1.80**	-0.269
mar	-0.010	-0.09	0.022	-0.092	-0.51	-0.017
univ	-0.050	-0.47	-0.024	-0.028	-0.12	0.032
city	0.217	1.62	0.213	0.015	0.08	0.007
inc	-0.002	-1.61	-0.001	-0.008	-1.36	-0.005
empl	0.358	1.94**	0.358	0.153	0.52	0.153
H1	1.083	3.89*	1.034	2.099	6.61*	1.988
H2	0.584	3.03*	0.567	1.362	4.81*	1.323
H3	1.281	5.86*	1.271	2.147	5.37*	2.126
H4	0.251	0.66	0.225	0.405	0.83	0.345
H5	0.883	7.04*	0.886	1.481	6.66*	1.488
limdy	0.084	6.16*	0.084	0.034	2.17*	0.033
pubins				0.447	1.82**	0.447
constant	-2.972	-4.54*		-4.516	-4.52*	
Theta	-0.267	-1.04		-0.388	-0.61	

Note: % Eff. refers to the marginal percentage effects for the corresponding variables. * and ** indicate significance at 5 and 10 percent, respectively.

Table 7. Two-Stage Method of Moments Estimation for Specialists Visits

	INSURED			UNINSURED		
Variables	**Coef.**	**t-stats**	**% Eff.**	**Coef.**	**t-stats**	**% Eff.**
sex	-0.574	-3.04*	-0.573	-0.257	-1.20	-0.256
age	-0.021	-0.34	-0.032	0.063	1.24	0.051
age2	0.320	0.46	0.530	-0.882	-1.52	-0.661
white	-0.053	-0.21	0.020	0.266	1.04	0.344
mar	-0.376	-1.23	-0.259	0.079	0.41	0.203
univ	0.274	1.80**	0.368	-0.056	-0.25	0.042
city	0.493	2.44*	0.481	0.242	1.06	0.229
inc	0.003	1.13	0.008	0.001	0.19	0.006
empl	0.183	0.79	0.183	-0.106	-0.36	-0.106
H1	1.569	4.85*	1.396	2.454	7.37*	2.271
H2	0.988	3.22*	0.927	1.566	4.32*	1.502
H3	1.210	3.46*	1.177	1.721	4.52*	1.686
H4	-0.024	-0.03	-0.116	0.541	0.88	0.443
H5	0.758	3.30*	0.770	1.157	3.91*	1.169
limdy	0.157	11.62*	0.156	0.074	4.67*	0.072
pubins				0.489	2.26*	0.489
constant	-3.215	-2.15*		-6.054	-4.63*	
Theta	-0.784	-2.29*		-0.687	-0.50	

Note: % Eff. refers to the marginal percentage effects for the corresponding variables. * and ** indicate significance at 5 and 10 percent, respectively.

Table 8. Two-Stage Method of Moments Estimation for Nights in the Hospital

	INSURED			UNINSURED		
Variables	**Coef.**	**t-stats**	**% Eff.**	**Coef.**	**t-stats**	**% Eff.**
sex	0.348	1.58	0.348	0.491	1.63	0.493
age	-0.095	-1.39	-0.095	-0.080	-1.12	-0.094
age2	1.157	1.38	1.157	0.946	1.09	1.206
white	0.761	2.36*	0.762	-0.220	-0.80	-0.128
mar	-0.447	-1.45	-0.446	0.220	0.71	0.365
univ	0.060	0.20	0.061	-0.009	-0.02	0.106
city	-0.197	-0.61	-0.197	-0.122	-0.43	-0.137
inc	-0.001	-0.22	-0.001	-0.010	-0.84	-0.004
empl	-0.446	-1.49	-0.446	-0.720	-1.80**	-0.720
H1	2.323	9.40*	2.322	1.788	5.83*	1.574
H2	1.354	5.45*	1.354	0.931	2.59*	0.856
H3	1.912	6.33*	1.912	1.017	2.78*	0.976
H4	2.085	3.93*	2.084	0.789	2.09*	0.675
H5	0.798	5.02*	0.798	0.530	2.20*	0.544
limdy	0.096	4.68*	0.096	0.091	4.71*	0.090
pubins				0.475	1.62	0.475
constant	-0.420	-0.36		-0.391	-0.23	
Theta	-0.003	-0.01		-0.844	-0.28	

Note: % Eff. refers to the marginal percentage effects for the corresponding variables. * and ** indicate significance at 5 and 10 percent, respectively.

Table 9. Two-Stage Method of Moments Estimation for Operations

	INSURED			UNINSURED		
Variables	**Coef.**	**t-stats**	**% Eff.**	**Coef.**	**t-stats**	**% Eff.**
sex	0.018	0.10	0.018	-1.033	-5.11*	-1.034
age	-0.096	-2.57*	-0.099	-0.189	-4.73*	-0.183
age2	0.996	2.02*	1.049	2.016	3.93*	1.914
white	0.000	0.00	0.019	0.236	1.12	0.200
mar	0.501	2.52*	0.530	0.402	2.79*	0.344
univ	0.136	0.93	0.160	-0.047	-0.29	-0.093
city	0.056	0.39	0.053	0.056	0.35	0.062
inc	0.000	-0.13	0.001	-0.004	-0.85	-0.006
empl	-0.366	-1.84**	-0.366	-0.564	-2.97*	-0.564
H1	0.679	2.71*	0.635	0.006	0.02	0.091
H2	0.687	3.62*	0.672	-0.230	-0.95	-0.200
H3	0.625	2.66*	0.617	-0.765	-1.47	-0.748
H4	0.404	0.89	0.380	-0.081	-0.25	-0.035
H5	0.352	2.76*	0.355	0.439	1.94**	0.433
limdy	0.127	11.11*	0.127	0.060	3.52*	0.061
pubins				0.732	4.44*	0.732
constant	-1.155	-1.71		1.014	1.70**	
Theta	-0.243	-0.85		0.258	1.37	

Note: % Eff. refers to the marginal percentage effects for the corresponding variables. * and ** indicate significance at 5 and 10 percent, respectively.