Opinion Artilce

iMedPub Journals
http://www.imedpub.com

Journal of Health & Medical Economics
ISSN 2471-9927

2016
Vol. 2 No. 1: 5

DOI: 10.21767/2471-9927.100011

Econometric Models for the Demand for Medical Care with Excess Zeros

Shih-Ti Yu[1] and Chii-Shyan Kuo[2]

1 Department of Quantitative Finance, College of Technology Management, National Tsing Hua University, Hsinchu, Taiwan
2 Department of Accountancy, College of Management, National Cheng Kung University, Tainan, Taiwan

Corresponding Author: Shih-Ti Yu

Department of Quantitative Finance, College of Technology Management, National Tsing Hua University, Hsinchu, Taiwan.

styu@mx.nthu.edu.tw

Tel: 886 3 5762130

Citation: Shih-Ti Yu. Econometric Models for the Demand for Medical Care with Excess Zeros. J Health Med Econ. 2015, 2:1.

Received: November 16, 2015; **Accepted:** November 18, 2015; **Published:** November 28, 2015

Introduction

Econometric models have played an important role in assessing the forces that have caused the increases in the demand for medical care. There is a large economic literature on the demand for medical care. Duan [1] point out the decision-making process of a patient. At the first stage, it is a person who decides whether to have a physician visit. Then the physician determines the intensity of treatment and therefore decides medical care expenditure. Therefore, the variables to represent the demand for medical care are usually studied: whether to have a physician visit, number of physician visits, and medical care expenditure. Medical care expenditure is most likely to increase as an increase in the number of physician visits. In the next session, we will briefly introduce some econometric models to fit the specific features of medical data.

Econometric Models

Whether to have a physician visit is a dichotomous dependent variable. For modeling the likelihood of visiting a physician, Probit and Logit models have been proposed for cross-sectional data, while Chamberlain's [2] fixed-effect logit estimator could be used for panel data. The fixed-effect logit model can lessen the problems of unobserved heterogeneity and omitted variable bias which would cause the cross-sectional data analysis invalid. While studying the number of physician visits, count data models such as Poisson and the negative binomial models are proposed, where the Poisson model assumes the equality between the mean and variance and the negative binomial model relaxes this assumption. However, the ability of these models to deal with zero values is quite limited because it depends on their mean and variance. In case of excessive number of zeros, alternative models such as zero-inflated Poisson and zero-inflated negative binomial models are used. Among these four models, the most generalized model is the zero-inflated negative binomial model because it can deal with the problems of excessive number of zeros and over-dispersion. The zero-inflated models separate the zeros into two different and independent groups: one group who never visit the physician, and the other group who will visit the physician if needed. In other words, the zero-inflated models assume some fraction of the zero-sample would permanently avoid the doctor. This statistical technique is widely used in microeconometric literature. For example, Schmidt and Witte [3] propose a “split population model” in which they assume that some fraction of the sample would never return to prison. Hsiao and Sun [4] propose a “one-sided survey response bias model” in which the response is biased in only one direction. While utilizing the zero-inflated models for the number of physician visits, the assumption of why a fraction of people avoid going to the doctor should be justified. In addition to a large proportion of the zero values, it is more complicated to analyze the medical care expenditure due to right-skewed distribution of the non-zero values, and self-selection bias. Usually using a logarithmic transformation of non-zero data, this can make to highly skewed distributions less skewed. However, Duan [1] find that the right tail of the distribution for non-zero values on the log-scale is still longer than the lognormal distribution. Self-selection bias problem arises when health plan enrollees can choose their plan based on their own health conditions. For example, using a data set from the 1994 Medicare Current Beneficiary Survey, Riley [5] find that Medicare beneficiaries enrolled in Health Maintenance Organization (HMO) are most likely healthier than those in Fee For Service (FFS). This implies that the choice between HMO and FFS is subject to a non-random sampling process. In dealing with censored data and self-selection bias, sample selection models with a bivariate normal distribution have been often suggested in the literature, for example, van de Ven and van Praag [6] while

Duan [1] and Manning, Duan, and Rogers [7] advocate a two-part model. In the first part of the two-part model, a probit is used to deal with the dichotomous outcomes of an individual whether to visit the physician and predict the probability of an individual conducting a physician visit. In the second part, the two-part model is to decide the level of medical care expenditure by using non-zero observations only. In contrast, the sample selection model takes into account those observations with zero values. There were vigorous debates on the choice between two part model and sample selection model. See Pohlmeier and Ulrich [8] and Leung and Yu [9] for a critical review of past studies. In resolving the debates, Leung and Yu [9] find that Heckman's [10] two-step estimator for the sample selection model is vulnerable to multicollineaity problem based on Monte Carlo studies. However, in empirical applications of health economics, the two-part models are generally favoured [11] based on the prediction ability of actual outcomes. The outperformance of the two-part model may be attributable to discarding the zero observations in the second stage estimation. On the other hand, the high degree of censoring due to excessive number of zeros may cause Heckman's two-step estimator to violate the bivariate normality assumption. Like the Poisson model and the negative binomial model, sample selection model may have a limited ability to deal with the problems of excessive number of zeros. A common approach would be to use sample selection model, in conjunction with the zero-inflated modeling technique. Recently sample selection models for panel data such as Kyriazidou's [12] estimator are free of the problems of excess zeros and multicollinearity [13]. The advantage of using the two-part model would disappear if empirical applications to panel data.

Conclusion

Estimation of a demand for medical care equations plays a vital role for helping public policymakers to control health care spending. To the best of our knowledge, the issue of excess zeros has previously been ignored in the literature on sample selection models. In this article, the argument that while using sample selection models the excessive number of zeros may lead to violations of the distributional assumption and affect their prediction accuracy. We encourage further study on modeling sample selection model to deal with excess zeros.

References

1 Duan N, Manning W, Morris C, Newhouse J (1983) A comparison of alternative models for the demand for medical care. Journal of Business & Economic Statistics 1: 115-126.

2 Chamberlain G (1984) Panel data. in: Griliches Z, Intriligator M (Eds), Handbook of Econometrics, North-Holland, Amsterdam, 2: 1247-1318.

3 Schmidt P, Witte A (1989) Predicting criminal recidivism using 'split population' survival time models. Journal of Econometrics 40: 141-159.

4 Hsiao C, Sun BH, (1999) Modeling survey response bias - with an analysis of the demand for an advanced electronic device. Journal of Econometrics 89: 15-39.

5 Riley G, Tudor C, Chiang Y, Ingber M (1996) Health status of medicare enrollees in HMO's and fee for service in 1994. Health Care Financing Review 17: 65-76.

6 Van de Ven WPMM, Gaag BV (1982) Health as an unobservable: a MIMIC - model of the demand for health care. Journal of Health Economics 1: 157-183.

7 Manning WG, Duan N, Rogers W (1987) Monte Carlo evidence on the choice between sample selection and two part models. Journal of Econometrics 35: 59-82.

8 Pohlmeier W, Ulrich V (1995) An econometric model of the two-part decisionmaking process in the demand for health care. The Journal of Human Resources 30: 339-361.

9 Leung SF, Yu S (1996) On the choice between sample selection and two-part models. Journal of Econometrics 72: 197-229.

10 Heckman JJ (1979) Sample selection bias as a specification error. Econometrica 47: 153-161.

11 Madden D (2008) Sample selection versus two-part models revisited: The case of female smoking and drinking. Journal of Health Economics 27: 300-307.

12 Kyriazidou E (1997) Estimation of a panel data sample selection model. Econometrica 65: 1335-1364.

13 Cameron C, Trivedi P, Milne F, Piggott J (1988) A microeconometric model of the demand for health care and health insurance in Australia. Review of Economics Studies 55: 85-106.