STOCKHOLM SCHOOL
OF ECONOMICS
THE ECONOMIC RESEARCH INSTITUTE

Equity in Health Care Utilisation: Further Tests Based on Hurdle Models and Swedish Micro Data

Ulf-G Gerdtham

Working Paper No. 116
June 1996

Working Paper Series
in Economics and Finance

1996-06-11

Equity in health care utilisation:

Further tests based on

Hurdle models and Swedish micro data*

Ulf-G Gerdtham, Centre for Health Economics, Stockholm School of Economics.

Abstract. This paper tests the null hypothesis of no horisontal inequity in delivery of health care by use of count data Hurdle models and Swedish micro data. It differs from most earlier work in three principal ways: First, the tests are carried out separately for physician and hospital care; second, the tests are carried out separately for the probability of seeking care and the amount of care received (given any use); and third, the tests are based on a model that includes several socioeconomic variables; e.g., income, education and size of community of residence. Most earlier work on testing for inequity has restricted attention to the bivariate relationship between income and health care utilisation (standardized for need=morbidity). The paper concludes that need, proxied by morbidity, has a significant positive effect on health care utilisation. Despite this, it rejects the hypothesis of no inequity because socioeconomic factors also have significant effects on utilisation; e.g., income and size of community of residence. Income has a significant positive effect on the probability of visiting a physician but not on the frequency of physician visits. Size of community of residence has a positive significant effect on the frequency of physician visits but not on the probability of visiting a physician. This latter finding is interpreted as evidence of departure from equity, as a result of supplier-induced demand.

JEL-Classification: C30, D63

Key words: Equity, Health Care, Count Data, Maximum Likelihood estimations.

***Acknowledgements:** This work was financially supported by the National Corporation of Swedish Pharmacies. I also thank for their comments three colleagues, Markus Asplund, Magnus Johannesson and Clas Rehnberg. Any errors are, of course, mine.

1. Introduction

Applied research on equity in the delivery of health care often concentrates attention on the horisontal version of the principle of distribution according to need, i.e., that individuals in equal need (in terms of morbidity) are treated the same (in terms of utilisation) irrespective of income (see Le Grand 1978; O'Donnell and Propper 1991; Wagstaff et al. 1991; Van Doorslaer and Wagstaff 1992; Doorslaer et al. 1993). In a recent contribution, Van Doorslaer and Wagstaff (1992) employed two methodologies to investigate the degree of income-related inequity in selected (mainly) European countries. The *first* methodology involves ranking individuals by income and then comparing the cumulative proportion of health care expenditure (standardized for differences in morbidity) across income groups. The *second* methodology controls for morbidity implementing regression analysis to test for significant income effects on health care received. This involves estimating a two-part model, where the first part relates to the determinants of an individual's decision to seek care, and the second to the determinants of the amount of care received given some utilisation. Controlling for age and gender, and with dummy variables for ill/not ill, using the responses on chronic illness and on self-reported health, the additional explanatory power provided by inclusion of income quintile dummies and interactions between income quintile and the other explanatory variables were examined. The results showed that inequity exists in most countries, but that there is no simple one-to-one correspondence between a country's delivery system and the degree to which individuals of equal morbidity have equal utilisation.

Despite significant progress in the methods used to measure and test income-related inequity, it is still not clear how the results should be interpreted. First, departure from income-related equity does not reveal where in the health care system inequity arises; e.g., is there inequity in physician or hospital care, or both? Second, departure from income-related equity does not reveal if it is the behaviour of the individual that changes with income, or if it is the behaviour of the health care provider that changes with individual income?[1] Third, departure from income-related equity does not even say if it is income that affects utilisation or if it is some other omitted variable(s) that are correlated with both income and utilisation. One such omitted variable that may be correlated with income is information. Individuals economically poorer may not be as well informed about health matters compared with those that are better-off, economically. This means that even those that are poorer have equal opportunities to receive, e.g., preventive care, it might be expected that the healthy poorer use less preventive care than the healthy richer because the take-up rate for the poorer is lower compared with the take-up rate for the better-off (Culyer et al. 1992). Another omitted variable is place of living. Poorer individuals tend to live in areas where health care resources are relatively scarce. This suggest that observed income-related inequity may reflect that the costs of seeking care are higher for individuals living in rural areas than in urbanized areas. It is also possible that the effects of omitted variables could be interpreted differently, depending on the stage of the decision making process; e.g., living in urbanized areas may represent an availability effect at the decision stage of seeking care (governed by the individual), but a supplier-induced demand effect at the decision of repeated visits and/or referrals (governed largely by the health care provider).

The aim of this paper is to test the null hypothesis of no horisontal inequity in the delivery of health care in Sweden (equity=equal treatment for equal need). To conduct the test, I first

[1] The latter may arise if the provider is paid on a fee-for-service basis for (richer) private patients but by salary or capitation for (poorer) public patients (Van Doorslaer and Wagstaff 1992).

developed a negative binomial (Negbin) Hurdle model to analyze physican visits and hospital care weeks, separately. The Negbin Hurdle model takes care of the discrete dependent variables (physician visits or hospital care weeks), and specifies the two stages of the decision making process of health care utilisation as different stochastic processes. Next, I used data from the Swedish Level of Living Survey and the developed model to test the null hypothesis of no inequity. The test is conducted separately for physician/hospital care and for both stages of decision making.[2] The test is also conducted separately for individuals reporting good health and fair/bad health.

The remainder of the paper is organized as follows: Section 2 presents the data and the variables selected; Section 3 describes the methods used in the paper; Section 4 reports the results, and Section 5 presents the conclusions.

2. Data and variable specification

The tests conducted are based on data from probability samples of the Swedish population, the Level of Living Survey (LNU) from 1991. The total sample consists of 7,856 individuals, between the ages 18-76 years in 1991. After correcting for missing values, the sample was reduced to 5,011 individuals in 1991. The survey contain data on morbidity, health care utilisation, and different socioeconomic variables. The available data on incomes (wages and transfers, including non-taxable transfers for both the repondent and the wife/husband/cohabitant) are of a high quality, because the LNU survey has been linked to national income tax statistics. The income data refer to the year before the interview, i.e.,

[2] Indeed, the tests of no inequity in Van Doorslaer and Wagstaff (1992) involve fitting the two-part model, but their tests are not carried out separately for physician and hospital care nor separately for each part of the two-

1990. Further details are given in Levnadsnivåundersökningen (1991) and Fritzell - Lundberg (1994).

The dependent variables in the models are the number of physician visits and the number of care weeks in hospital. Each subject was asked how many times he/she had visited a physician in the last twelve months and how many weeks he/she had spent in hospital.

The explanatory variables included comprise, in addition to a constant denoted ONE, 19 variables that are categorized according to whether they reflect need or socioeconomic factors.

A. Need (=morbidity) factors: I first included two 0-1 dummies for individuals reporting bad health (BADHLTH=1, zero otherwise) and fair health (FAIRHLTH=1, zero otherwise). The reference category is individuals reporting good health. Each subject in the survey was asked to judge his/her own present health condition on a 3-point scale from good to bad. To further account for morbidity, I included two 0-1 dummies; the first is for individuals reporting at least one chronic illness, but state that they are able to run, go up and down stairs, and turn on the water-tap (NLCHRONIC=1, zero otherwise); i.e., not limiting chronic illnesses. The other dummy is for individuals reporting at least one chronic illness, but state that they are ***not*** able to run or walk up steps or turn on the water-tap (LCHRONIC=1, zero otherwise); i.e. limiting chronic illnesses. Each subject in the survey was asked whether he/she has been suffering severely, mildly or not at all concerning 44 illness conditions during the past 12 months. Eleven of these 44 illness conditions have been defined as chronic conditions (i.e., cancer, diabetes, elevated blood pressure, heart attack/myocardial infarction, heart failure, chronic bronchitis, gastric ulcer/duodenal ulcer, gall bladder, kidney disease, rash or excema, or joints

---

part model.

ache). If the individual reports that he suffers *severly* from at least one of these conditions, then he is classified as having a chronic illness. Further, gender and age are included in the model, which may capture imperfect measurements of morbidity for individuals of different ages or sex, e.g., obstetrical charges. Gender is represented by a 0-1 dummy for male (MALE=1, zero otherwise), which means that females are the reference category. Rather than impose a functional form for age on utilisation, I have conservatively used three 0-1 dummies for age groups; i.e., individuals in the age ranges 35-44 years (AGE2=1, zero otherwise), 45-64 years (AGE3=1, zero otherwise) and 65-76 years (AGE4=1, zero otherwise). The reference category is the age group 18-34 years.

B. Socioeconomic factors: Again to avoid the problem of functional form, I include four 0-1 income dummies: INC2=1 if the disposable household income corresponds to the second quintile, zero otherwise; INC3=1 if the disposable household income corresponds to the third quintile, zero otherwise, INC4=1 if the disposable household income corresponds to the fourth quintile, zero otherwise; and INC5=1 if the disposable household income corresponds to the fifth quintile, zero otherwise. The reference category is individuals that belong to the first (poorest) quintile of disposable household income. To deflate household income to per person levels, I include number of persons in the houshold (FAMSIZE). I further use a 0-1 dummy for employment status (WORKING = 1 if currently working, zero otherwise). The argument is that individuals who are employed may incur a larger ”time price” of going to the physician than the non-employed (Phelps and Newhouse 1974; Cauley 1987). To measure life satisfaction a 0-1 dummy variable is included: (DSATIS=1 if the daily life of the individual is a source of personal satisfaction, zero otherwise). Also I include EDUCATION, measured in terms of years of formal schooling. The a priori expectation of education, which eventually is correlated with medical knowledge, is ambiguous (Grossman 1972, Muurinen 1982, Wagstaff

1986,1993). I include marital status (MARRIED=1 if married, zero otherwise), which may have a negative effect on use, because single persons generally use more health care (Feldstein 1979). Finally, two covariates on community size are included as 0-1 dummies for individual living in big cities (BIGCITY=1 if the individual lives in Stockholm, Gothenburg or Malmö, zero otherwise) and for individuals living in small cities (SMALLCITY=1 if the individual lives in cities larger than 30,000 inhabitants, zero otherwise).

The frequency distribution of physician visits and care weeks in hospital are presented in Table 1, and the descriptive statistics for the estimation sample are presented in Table 2. Some of the characteristics of the raw data on physician visits and care weeks are as follows: 38.8% of 5,011 respondents interviewed had zero visits, 21.9% had one visit and the remainder had up to a maximum of 50 visits; 91.3% had zero care weeks in hospital, 5.0% had one visit and the remainder had up to a maximum of 52 care weeks in hospital.

(Table 1 about here)

(Table 2 about here)

3. Estimation methods

Utilisation of health care has two characteristics that are important in selecting an estimation method: one characteristic is that the distribution of the number of physician visits and hospital care weeks can take only nonnegative integer values, which means that some individuals had no physician visits and no care weeks during the survey, whereas others had single or multiple visits and care weeks. This calls for the application of count data models (Cameron et al. 1986, 1988), and the Poisson model may represent a natural starting point for estimating number of physician visits and care weeks. Assuming a random variable y which can take only nonnegative integer values, the probability that y will occur N times is (with N=0,1,2,...) as follows:

$$\text{Prob}(y_i=N) = (e^{-\lambda}\lambda_i^N)/N! \quad \text{for } N=0,1,...,\infty \qquad (1)$$

with y_i being the count of the relationships for the physician visits and care weeks in hospital of the i^{th} individual. To incorporate exogeneous variables, lambda can be made a function of the covariates:

$$\lambda_i = \exp(\Sigma b_j X_{ji}), \qquad (2)$$

where b‘s are the coefficients. X’s are the covariates (with X_1 set to one), j indicates the j^{th} variable, and i is the i^{th} individual. The exponential function ensures non-negativity.

The Poisson model is not unreasonable given the data illustrated in Table 1, where the distribution of physician visits and care weeks resemble a Poisson distribution. Furthermore, the Poisson model is easy to interpret because it can be seen as average proportionate changes in $E(y_i | x_i)$ for a unit change in x_i, i.e.:

$$\frac{\partial E(y_i | x_i)}{\partial x_{ij}} = \beta_j E(y_i | x_i)$$

In applied work, however, the Poisson model is restrictive in that it assumes that the mean is equal to the variance. If this restriction is violated, the coefficients are consistent but their standard errors are not (see e.g., Cameron and Trivedi 1990). In most empirical data concerning economic behaviour, this mean/variance assumption is rarely valid. When we look at our data in Table 1, we can see that the sample mean for the number of physician visits is 1.95 and the sample variance is 11.97, indicating considerable overdispersion in the raw data. This overdispersion might be the result of state dependence; i.e., the probability to visit a physician today might be dependent on the visit of the previous day. The Poisson distribution can be viewed as the result of a process of events (physician visits or care weeks), the timing of which is independently exponentially distributed. Overdispersion might thus be the result of a violation of the assumption of independence (Amemiya 1985, p. 436, Cameron and Trivedi 1986, p. 31). Because my available data only contains the total number of physician visits or total number of hospital care weeks during a year, and not the single spells of visits and care weeks, I'm unable to estimate models with state dependence. Hence, I must model the overdispersion displayed in the data in some alternative way. Overdispersion in count data models has been discussed extensively (see Hausman et al. 1984, Cameron and Trivedi, 1986, McCullagh and Nelder, 1983), and formal tests of overdispersion have been devised by Cameron and Trivedi (1990).

A generalisation of the Poisson model, as suggested by McCullagh and Nelder (1983,p. 194), is the negative binomial model (Negbin) in which the variance/mean ratio is linear to the mean. The Negbin model has all the advantages of the Poisson model but without its constraints; i.e. the Negbin model captures the discrete, censored and overdispersion properties in the data. With the Negbin distribution equation (2) becomes:

$$\lambda_i = \exp(\Sigma b_j X_{ji})\exp(e_i), \qquad (3)$$

where λ_i is no longer determined but is itself a random variable. Because e_i is unobserved, it is integrated out of the expression by specifying a gamma distribution for the error term. Estimation of the count data is done by the maximum likelihood (ML) procedure, which gives us both consistent and efficient estimates from the Negbin model, given overdispersed discrete data.

The second characteristic of health care utilisation is the two-part decision making process (Manning et al. 1981), where the first part relates to the patient who decides whether to contact the physician (contact decision), and the second to the decision about repeated visits and/or referrals, which is determined largely by the preferences of the physician (frequency decision; see Duan et al. 1983, Zweifel 1992, pp. 25). In standard Poisson and Negbin models, this two-part character is ignored which may lead to inconsistent parameter estimates and hence to misinterpretation. The econometric specification used in this study is based on the Hurdle model for count data as proposed by Mullahy (1986). The model decomposed the observed count (y_i) into two observed random variables. These two variables are, e.g., whether or not an individual had visited a physician or not: "y> 0" and how many physician visits an individual had done if he/she had visited a physician at all: "$y_i \mid y_{ui} > 0$". By separating the decision of any visit from the frequency of visits, it may be possible to assess whether income, e.g., has its effect largely through the contact decision or the frequency decision in physician care.

The decision to hospitalize, however, cannot be interpreted in the same way as the decision to see a physician insofar as the first stage of decision making in hospital care is likely to be determined by a physician. On the other hand, this does not imply that the two stages of the decision making process are identical for hospital care. It is not implausible to envisage that the decision to hospitalize is controlled mainly by the general practioner or the practicing specialist, whereas the frequency (and treatment intensity) decision could be under the control of a hospital physician (Zweifel 1985). For this reason, I use the Hurdle model to test the possibility that the decision of any use of hospital care is different from the frequency decision. A further problem is that I have no information about the number of hospital care days only hospital care weeks. This means that I have to assume that the decision of admission to a hospital is made on the basis of a week and not on the basis of a day; i.e., the probability of any hospital care is modelled as the probability of any hospital care week.[3]

In the estimation of the Hurdle model, I use two equations to model the distribution of physician visits and hospital care weeks, respectively. In the first equation, I specify a binary (zero/positive) logit model for the probability that the individual will visit a physician and will have a care week, i.e.

$$\text{Prob(Visits or care weeks>0)} = 1/[1+ \exp(-X_i*a], \qquad (4)$$

where X_i is a row vector of k given individual characteristics (e.g., gender, age) and a is a set of parameters to be estimated.

[3] If this assumption is wrong, then one may expect that the standard Poisson and Negbin model could not be rejected against the Hurdle model.

In the second equation, I specify a truncated-at-zero Poisson or Negbin model to model the number of physician visits and care weeks attended by the individual, i.e:

$$(\lambda_i \mid \lambda_i > 0) = \exp(\Sigma b_j X_{ji})\exp(e_i), \qquad (5)$$

where X_i is a row vector of k given individual characteristics (e.g., gender, age), b is a set of parameters to be estimated and e_i is the error, which is conditionally independent of X.

A wide range of specification tests on the Negbin Hurdle model is carried out. To discriminate between the Hurdle model and standard Poisson/Negbin models, I use likelihood ratio and Wald tests.[4] To test for over- and underdispersion in the data, I use a test devised by Cameron and Trivedi (1990, p. 353; μ, μ^2). The test is based on the following hypothesis:

H_0: $\quad Var(y_i) = \mu_i$

H_1: $\quad Var(y_i) = \mu_i + \alpha * g(\mu_i)$

[4] I have also experimented with some alternative count data specifications, i.e. the zero-inflated Poisson and Negbin (ZIP) models (see Lambert 1992 and Greene 1995), and different sample selection models for count data (see Greene 1995). However, the zero-inflated Poisson model of physician visits failed to converge when the splitting model was a function of all the regressors. Furthermore, the Vuong test statistic (Vuong 1989) testing the standard Negbin model of physician visits against the zero-inflated Negbin model gave an inconclusive result, i.e., it was not possible to reject any of these models against each other because the Vuong statistic was neither higher than 1.96 nor lower than -1.96. However, the estimates of the zero-inflated Negbin model are similar to the standard Negbin. The zero-inflated Poisson model of hospital care weeks also failed to converge but the zero-inflated Negbin seemed to be superior to the standard Negbin, i.e. the Vuong test was 2.79 (higher than 1.96). The results, however, were again similar to the standard Negbin. The sample selection model for count data described by Greene(1995) failed to converge for both the number of physician visits and hospital care weeks, irrespective of estimating by use of the two-step limited information ML or the full information ML, and even if all insignificant variables were omitted at each stage of estimation.

where μ_i equals the mean and $g(\mu_i)$ some function of the mean. In the test they propose, their optimal regression is carried out by testing the significance of the single coefficient in the linear OLS regression of:

$$(\sqrt{2*\mu_i})^{-1} \text{ x } [(y_i - \mu_i)^2 - y_i] = (\sqrt{2*\mu_i})^{-1} \text{ x } g(\mu_i) * \alpha + \varepsilon_i,$$

and they suggested two possibilities for $g(\mu_i)$, i.e. μ_i and μ_i^2.

In addition, to check for robustness of the results, I re-estimate the models without observations in the sample reporting more than 10 physician visits or care weeks. Also, I re-estimate the models, excluding variables with insignificant effects, that may increase the available sample for estimation.[5] These additional tests of robustness are discussed where they influence the overall conclusions.

4. Results

Tables 3 and 4 present the estimation results with the covariates included for number of physician visits and hospital care weeks, respectively. In both tables, the first column gives the ML estimated effects of the restricted (standard) Poisson model. The second column presents the ML estimated effects of the restricted (standard) Negbin model.[6] The third and fourth columns present the estimated effects of the unrestricted Negbin Hurdle models that are obtained by separate ML estimation of the binary logit model estimated over the entire

[5] Detailed estimation results from these additional tests are not provided in this paper, but are available from the authors on request.

sample, and a truncated-at-zero Negbin model estimated on the sample having positive realizations of physician visits or care weeks in hospital (see Mullahy 1986).

Tables 5 and 6 report subsample estimations of physician and hospital care for individuals reporting good health (healthy) and fair/bad health (unhealthy), respectively.

[6] Referring to Cameron and Trivedi (1986, pp. 32-33); this is their model Negbin II.

A. Estimating number of physician visits

In Table 3, the LR test statistic for testing the restricted Poisson model against the restricted Negbin model ($\chi 2$ (1)) is 5286.604 (2 $\times$ (11658.04-9014.738)), which is highly significant at the 1% level. The corresponding Wald statistic (N (0,1) is 33.2 (=1.101/0.0331) and, again, the Poisson model is rejected. The estimated overdispersion parameter (α=1.101) of the Negbin model is positive, indicating overdispersion in the data. Next, the restricted Negbin model is tested against the Hurdle Negbin model. The resulting LR test statistic ($\chi 2$ (1)) is 114.384 (2 $\times$ (9014.738-(3089.786+5867.760)), which is highly significant at the 1% level; i.e., the restricted Negbin model is rejected against the Negbin Hurdle specification. This points to important differences between the two decision making processes of physician visits, and that both Poisson and Negbin restricted models result in inconsistent estimates, producing serious misinterpretation. I also tested the unrestricted Poisson Hurdle model against the Negbin Hurdle model. The resulting LR statistic ($\chi 2$ (1)) is 1389.991 (and the corresponding Wald statistic (N (0,1) is 11.3 (=1.639/0.145)), indicating that the Poisson Hurdle model must be rejected at the 1% level against the Negbin Hurdle. The two overdispersion tests ($\mu;\mu^2$) are clearly significant, for both the standard Poisson model and the truncated-at-zero Poisson model.

Because all restricted representations of the Hurdle model have been rejected, I further concentrate my discussion on parameter estimates of the Negbin Hurdle model (binary logit and the truncated-at-zero negative binomial model).

Moving on to the results:

** Need factors (=morbidity).* Not surprisingly, the contact and frequency decisions are clearly responsive to need, proxied by morbidity. The estimated effects pertaining to the dummies for self-reported health (BADHLTH=1, FAIRHLTH=1) and limiting and unlimiting chronic illnesses (LCHRONIC=1, NLCHRONIC=1) are positively significant at the 1% level. The effect of bad health also tends to be larger than the corresponding effect of fair health. The coefficients of the age dummies (AGE2-4) are generally negatively significant on both decision stages, indicating that individuals older than 34 years, *ceteris paribus*, have less physician visits than the reference category, i.e., people in the age range 18-34 years. Further, investigations suggest that the effect of AGE4 is negatively significant for both healthy and unhealthy people on the frequency decision but not on the contact decision, and that this effect on the frequency decision, is higher for unhealthy compared with healthy people.[7] The estimated effect of the male dummy (MALE) is negatively significant at both stages of decision making.

Considering next the covariates reflecting socioeconomic factors:

** Socioeconomic factors.* The dummy for employed people (WORKING) has a positive significant effect on the contact decision, which is not consistent with a priori expections; it is, however, not significant on the frequency decision. Additional subsample estimations suggest that the contact decision is only responsive to WORKING for healthy people. One, admittedly

[7] One possible explanation for the negative effect of AGE4 on the frequency decision of physician care is that older people, *ceteris paribus*, are reffered more often to a hospital. To test this, I estimated a binary logit model

ad hoc, explanation is that the healthy employed use more preventive care through their work than the healthy unemployed. Number of people in the household (FAMSIZE) is negatively significant on both decision stages. The subsample estimations indicate further that FAMSIZE is negative at both decision stages for healthy people but not significant at any decision stage for unhealthy people. The estimated effect of the dummy for those married (MARRIED) is negatively significant on the frequency decision but not on the contact decision. Number of years of schooling (EDUCYEARS) is not significant on any decision stage. The estimated effects of the four income quintile dummies (INC2-INC5) are all positive, and significant at the 10% level in two cases (INC3-4), on the contact decision, but no income dummy is significant on the frequency decision.[8] The subsample estimations show further that all four income dummies are positively significant at the 5% level at the contact decision stage for unhealthy people but in no case significant for healthy people. These results indicate that the probability of visiting a physician is higher for the better-off compared with the worse-off or, more precisely, that the probability of visiting a physician is higher for the unhealthy better-off compared with the unhealthy worse-off and that it appears to be no difference in the corresponding probability between the healthy better-off and the healthy worse-off. The estimated effect of the dummy for individuals living in big cities (Stockholm, Gothenburg or Malmö; BIGCITY) is positively significant at the 1% level on the frequency decision, but fails to reach significance on the contact decision. The additional estimations presented in Table 5 show further that BIGCITY is positively significant on the frequency decision for unhealthy people and close to significance for healthy people at the 10% level (t-value=1.63) but clearly

---

on the probability of having a hospital care week (given any visit to a physician) and found that the estimated effect of AGE4 in this model was positive (0.178), but clearly not significant (z=1.18).

[8] I also re-estimated a parsimonious representation including income and the square of income on the probability of visiting a physician and found that the estimated effects of these variables were positively and negatively significant, repsectively, at the 5% level. This indicates that the probability of visiting a physician increases with income at a decreasing rate.

not significant on the contact decision both for unhealthy and healthy people.[9] I interpret this finding as evidence for supplier-induced demand because the probability for having a physician visit (the contact decision) is not higher for individuals living in the big cities of Sweden, but given any visit it appears that the number of visits (the frequency decision) are higher for people living in big cities.

[9] The logit model on the probability of having a hospital care week, given any physician visit, (see footnote 7) shows a negative significant effect of BIGCITY (-0.136) at the 10% level. This indicates that the mix of health care utilisation is different for people living in big citites compared with people living in smaller cities. However, I also estimated the same logit model separated for unhealthy people. These subsample estimations show that the effect of BIGCITY is clearly not significant for unhealthy people. This implies supplier-induced demand because the effect of BIGCITY is positively significant on the frequency decision of physician care for unhealthy people but neither significant on the contact decision nor significantly negative on the the probability of having a hospital care week.

B. Estimating number of care weeks

It appears that all nested models of the Negbin Hurdle model for hospital care weeks can be rejected. The LR statistic for testing the restricted Negbin model against the Negbin Hurdle model ($\chi 2$ (1)) is 126.9 ($2 \times$ (2108.610-(1332.789+712.371)), which is highly significant, i.e. the restricted Negbin is rejected. The LR statistic for testing the unrestricted Poisson Hurdle model against the corresponding Negbin Hurdle model ($\chi 2$ (1)) was 519.009, which is significant at the 1% level; i.e. the Poisson Hurdle is rejected against the Negbin Hurdle. However, this result contradicts the corresponding Wald statistic (N (0,1) which is clearly insignificant 0.002 (=5098.5/2549250). I assume that this can be explained in that a very small part of the sample has more than 0 hospital weeks, namely, 436 individuals, and only 185 individuals have more than 1 care week in hospital. The two overdispersion tests ($\mu;\mu^2$) are highly significant for both the standard Poisson model and the truncated-at-zero Poisson model.

Turning to the Hurdle model results:

* *Need factors (=morbidity).* The probability of having a care week and frequency of care weeks in hospital are responsive to BADHLTH and FAIRHLT, and the effect of BADHLTH is nearly twice as high as the effect of FAIRHLTH on both decision stages. The estimated effects of the dummies for LCHRONIC and NLCHRONIC are also positively significant on the probability of having a care week but not on the frequency of care weeks. The results showed further that the dummies for age are generally not significant at any decision stage. One exception is AGE3, representing people in the age range 45-64 years, which is negatively

significant on the probability of having a care week but positively significant on the frequency of care weeks. The subsample estimations showed further that AGE3 is only significant on the probability of having a care week for healthy people. The estimated effect of MALE is negatively significant on the probability of having a care week (10% level) but insignificant with respect to the frequency of care weeks. The subsample estimations showed further that the effect of MALE is not significant on the frequency decision for both healthy and unhealthy people, but negatively significant on the probability of having a care week for healthy people and positively significant at the 10% level for unhealthy people. This means that the probability of having a care week tends to be lower for healthy males compared with healthy females, but higher for unhealthy males compared with unhealthy females.

Considering next the results concerning the socioeconomic variables.

* *Socioeconomic factors.* The effects of MARRIED, FAMSIZE, WORKING and EDUCYEAR are not significant on any decision stage. The estimated effects of the four income quintile dummies appear generally to be positively significant on the probability of having a care week, whereas they are negatively significant on the frequency of care weeks.[10] The subsample estimations for healthy and unhealthy people show the same result concerning the income effects, and that the estimated effects on the probability of having a care week generally are higher for the subsample of unhealthy people compared with the subsample of healthy people; e.g., the estimated effect of INC5 (highest income quintile) for the unhealthy is twice as high as for the healthy people. The estimated effect of BIGCITY is negatively significant on both decision stages. However, the subsample estimations show that BIGCITY

is only negatively significant at the 10% level on the probability of having a care week for healthy people.[11]

[10] The negative effect of the income quintile dummies on the frequency of care weeks are highly sensitive to the presence of outliers In a re-estimation of the truncated-at-zero Negbin model without individuals with more than 10 hospital care weeks, i.e. 26 individuals, I found that all four income dummies were clearly insignificant.
[11] The estimated effects of BIGCITY is highly sensitive to the presence of outliers and is clearly insignificant on the frequency decision if one exclude individuals with more than 10 hospital care weeks.

5. Conclusions

On the basis of data from the Swedish Level of Living Survey of 1991 and negative binomial Hurdle models of health care utilisation, I have tested the null hypothesis of no inequity (equity=equal treatment for equal need) in the delivery of health care in Sweden. Unlike most past studies, the tests were carried out separately with respect to: (a) physician/hospital care, and (b) contact/frequency decision of use. Moreover, the tests were based on models that included socioeconomic variables other than income as regressors (e.g., dummy variables for employment, education and size of community of residence). This exercise was an attempt to establish if, where, why and how inequity occurs.

Based on the findings of the study some tentative conclusions can be drawn.

First, I reject the null hypothesis of no horisontal inequity in the delivery of health care in Sweden for both physician and hospital care. Though the effect of need, proxied by morbidity, on utilisation of health care was clearly significant (consistent with vertical equity), it is rather evident that socioeconomic factors, such as income, education, place of residence, household size, marital status and professional status were, to some extent, significant for either physician or hospital care, or both.

Second, the contact decision of physician care tends to be positively associated to individual income (at a decreasing rate), but income was not significant on the frequency of physician visits. I also re-estimated the model of physician visits separately for healthy (reporting good health) and unhealthy (reporting fair or bad health) people. The results of these subsample estimations indicated that the effect of income (all four income quintile dummies) was positively significant at the 5% level on the contact decision for unhealthy people, but not for healthy people and never significant on the frequency decision. I interpret these findings as evidence that income affects the individual's decision to visit a physician but not the decision of the physician, at least not in a favourable direction for the better-off.

Third, the income effect on the probability of having a care week was generally positively significant. I found generally the same results in the subsample estimations for healthy and unhealthy people. The estimated income effect on the probability of having a care week also tended to be higher for the subsample of unhealthy people compared with the corresponding effect for healthy people. This indicates that the physician's (general prationer or practicing specialist) admission decision of a patient to a hospital is partly somehow affected by the income of the patient.

Fourth, the income effect on the frequency of care weeks was in general negatively significant. Further investigations showed that this clearly holds for the subsample of healthy people but not as clearly for the subsample of unhealthy people. This is because only one of the four income dummies was negatively significant at the 10% level in that latter estimation, and the dummies representing the two highest income quintiles were clearly not significant. These results may be interpreted to mean that the economically poorer patients are treated favourably by the provider (hospital physician) in relation to the better-off in hospital care. A plausible alternative explanation is that the negative income effects reflect income-related differences in morbidity, which have not been entirely captured by self-reported morbidity measures used in this study. I further found that the results of the income effects on the frequency decision in hospital care were highly sensitive to outliers. When individuals with more than 10 hospital care weeks were excluded from the sample, there were no significant income effects on the frequency decision in hospital care.

Fifth, the estimated effect of family size was negatively significant on both the contact and the frequency decision concerning physician care. In the additional subsample estimations, I found that the effect of family size was negatively significant on the contact and frequency decision at the 10% level for healthy people only. This indicates that the probability of a visit to a physician tends to be lower, *ceteris paribus*, for healthy people living in larger families compared with healthy people in smaller families, but that there is no corresponding difference between unhealthy people living in large families and unhealthy people living in smaller families.

Sixth, the effect of the dummy for employed people was positively significant on the contact decision but not on the frequency decision of physician visits. I also re-estimated the model of physician visits separately for healthy (reporting good health) and unhealthy (reporting fair or bad health) people. The results of these subsample estimations indicated that the dummy for the employed was not significant on any stage of decision making for the unhealthy, but a positive significant effect was observed on the contact decision for the healthy. An ad hoc explanation of the results regarding the effect of the employed is that the healthy employed use more preventive care through their work compared with the healthy unemployed.

Seventh, the estimated effect of the dummy for people living in big cities (Stockholm, Gothenburg and Malmö) was positively significant on the frequency of physician visits but not on the contact decision. I interpret this as evidence for inequity in physician care as a result of supplier-induced demand. In the subsample estimations for healthy and unhealthy people, I found that the dummy for people living in big cities was positively significant on the frequency decision for both healthy and unhealthy people, but it was never significant on the contact decision.

References

Amemiya, T., 1985, Advanced Econometrics, Harvard University Press, Cambridge.

Cameron, A.C. and Trivedi, P.K., 1986, Econometric models based on count data: comparisons and applications of some estimators and tests, Journal of Applied Econometrics 1, 29-53.

Cameron, A.C., Trivedi, P.K., and Milne, F., 1988, A microeconometric model of the demand for health care and health insurance in Australia, Review of Economic Studies LV, 85-106.

Cameron, A.C. and Trivedi, P.K., 1990, Regression based tests of overdispersion in the Poisson model, Journal of Econometrics 46, 255-274.

Culyer, A.J., Doorslaer, van, E., and Wagstaff, A., 1992, Utilisation as a measure of equity by Mooney, Hall, Donaldson and Gerard, Journal of Health Economics 11, 93-98.

Doorslaer, van, E. and Wagstaff, A., 1992, Equity in the delivery of health care: some international comparisons, Journal of Health Economics 11, 389-411.

Doorslaer, van, E. et al., 1993, Equity in the finance and delivery of health care: an international perspective, Oxford University Press, Oxford.

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

Feldstein, P.J., 1979, Health care economics, John Wiley & Sons, New York.

Fritzell, J., Lundberg, O., (eds.) (1994), Vardagens villkor. Levnadsförhållanden i Sverige under tre decennier. Brombergs förlag. (The conditions of Daily Living)

Greene, W., 1993, Econometric analysis, 2nd Edition, Macmillan, New York.

Greene, W., 1995, LIMDEP Version 7.0 User's Manual, Econometric Software, Inc, Australia.

Grossman, M., 1972, The demand for health: a theoretical and empirical investigation, New York, Columbia University Press.

Hausman, J., Hall, B.H., and Griliches, Z., 1984, Econometric models for count data with an application to the patient: R&D relationship, Econometrica 52, 909-938.

Lambert, D., 1992, Zero-inflated Poisson regression, with an application to defects in manufacturing, Technometrics 34, 1-14.

Le Grand, J. 1978, The distribution of public expenditure: the case of health care, Economica 45, 125-142.

Manning, W.G., et al., 1981, A two-part model of the demand for medical care: preliminary results from the Health Insurance Study. In Health, Economics, and Health Economics, ed. Jacques van der Gaag and Morris Perlman, 103-123, Amsterdam: North-Holland.

McCullagh, P. and Nelder, J.A., 1983, Generalised Linear Models, Chapman and Hall, London.

Mullahy, J., 1986, Specification and testing of some modified count data models, Journal of Econometrics 33, 341-365.

Muurinen, J-M., 1982, Demand for health: A generalised Grossman model, Journal of Health Economics 1, 3-28.

O'Donnell, O., and Propper, C., 1991, Equity and the distribution of UK National Health Service resources, Journal of Health Economics 10, 1-20.

Vuong, Q., 1989, Likelihood ratio tests for model selection and non-nested hypotheses, Econometrica 57, 307-334.

Wagstaff, A., 1986, The demand for health: some new empirical evidence, Journal of Health Economics 5, 195-233.

Wagstaff, A., Doorslaer, E. van, and Paci, P, 1991, On the measurement of horisontal equity in the delivery of health care, Journal of Health Economics 10, 169-205.

Wagstaff, A., 1993, The demand for health: an empirical reformulation of the Grossman model, Health Economics 2, 189-198.

Zweifel, P, 1985, Individual choice in social health insurance: a curb on inflation in the health care sector? Evidence from Switzerland, in: Terny, G. and Culyer, A.J. (eds), Public Finance and Social Policy, Detroit: Wayne State University Press, 303-318.

Zweifel, P, 1992, Bonus options in health insurance, Kluwer Academic Publishers, London.

TABLE 1: SAMPLE FREQUENCY DISTRIBUTION OF THE NUMBER OF PHYSICIAN VISITS (NUMBER OF OBSERVATIONS=5,011)			
VISITS	DISTRIBUTION OF VISITS (%)	CARE WEEKS	DISTRIBUTION OF CARE WEEKS (%)
0	38.8	0	91.3
1	21.9	1	5.0
2	14.6	2	1.5
3	7.6	3	.5
4	5.0	4	.5
5	3.7	5	.2
6	2.3	6	.1
7	1.1	7	.1
8	.7	8	.2
9	.1	10	.1
10	1.6	11	.0
11	.0	12	.1
12	.8	14	.0
13	.0	16	.1
14	.0	17	.0
15	.6	19	.0
16	.1	20	.1
18	.0	22	.0
20	.5	27	.0
21	.0	30	.0
24	.0	31	.0
25	.1	40	.0
26	.0	52	.1
28	.0		
29	.0		
30	.1		
31	.0		
32	.0		
40	.0		
41	.0		
44	.0		
45	.0		
50	.1		
y (mean)	**2.06**	**y (mean)**	**0.30**
s^2y (variance)	**13.84**	**s^2y (variance)**	**5.11**
Variance/mean	**6.72**	**Variance/mean**	**17.03**

TABLE 2: SAMPLE DESCRIPTIVE STATISTICS
(N=NUMBER OF OBSERVATIONS).

VARIABLE	MEAN	STD DEV	MINIMUM	MAXIMUM	N
VISITS	2.06	3.72	.00	50.00	5,011
HOSPWEEK	.30	2.26	.00	52.00	5,011
MALES	.51	.50	.00	1.00	5,011
AGE2	.39	.49	.00	1.00	5,011
AGE3	.32	.47	.00	1.00	5,011
AGE4	.15	.36	.00	1.00	5,011
BADHLTH	.04	.20	.00	1.00	5,011
FAIRHLTH	.18	.39	.00	1.00	5,011
NLCHRONIC	.08	.27	.00	1.00	5,011
LCHRONIC	.08	.28	.00	1.00	5,011
MARRIED	.66	.47	.00	1.00	5,011
FAMSIZE	2.62	1.34	1.00	10.00	5,011
WORKING	.56	.50	.00	1.00	5,011
DSATIS	.59	.49	.00	1.00	5,011
EDUCYEAR	10.90	3.39	0	35	5,011
INC2	.2	.4	.00	1.00	5,011
INC3	.2	.4	.00	1.00	5,011
INC4	.2	.4	.00	1.00	5,011
INC5	.2	.4	.00	1.00	5,011
BIGCITY	.27	.45	.00	1.00	5,011
SMALLCITY	.22	.42	.00	1.00	5,011

TABLE 3: ESTIMATION RESULTS: DEPENDENT VARIABLE : FREQUENCY OF PHYSICIAN VISITS (COVARIATES INCLUDED).

	RESTRICTED				NEGBIN HURDLE MODELS			
VARIABLE	**POISSON**		**NEGBIN**		**BINARY LOGIT**		**TRUNCATED NEGBIN**	
	b[a]	z=b/s.e[b]	b	z=b/s.e	b	z=b/s.e	b	z=b/s.e
ONE	0.359***	5.98	0.435***	4.89	0.291	1.62	0.315**	2.34
MALE	-0.175***	-8.48	-0.207***	-6.06	-0.420***	-6.59	-0.931E-01*	-1.95
AGE2	-0.172E-03	-0.01	-0.333E-01	-0.57	-0.229**	-2.07	0.743E-01	0.91
AGE3	-0.235***	-5.72	-0.222***	-3.26	-0.340***	-2.74	-0.183*	-1.89
AGE4	-0.252***	-5.34	-0.205***	-2.73	0.102E-01	0.07	-0.320***	-3.12
BADHLTH	1.156***	30.55	1.156***	13.34	1.578***	6.07	1.154***	10.56
FAIRHLTH	0.874***	34.60	0.877***	20.61	1.010***	10.26	0.858***	14.77
LCHRONIC	0.755***	25.18	0.770***	13.13	1.057***	6.03	0.748***	9.69
NLCHRONIC	0.640***	21.08	0.731***	11.22	1.203***	8.32	0.610***	7.45
MARRIED	-0.917E-01***	-2.62	-0.106	-1.60	0.841E-01	0.76	-0.216**	-2.28
FAMSIZE	-0.538E-01***	-5.37	-0.496E-01***	-3.56	-0.676E-01**	-2.36	-0.405E-01**	-2.04
WORKING	0.526E-01**	2.100	0.296E-01	0.79	0.179**	2.35	-0.438E-01	-0.83
EDUCYEAR	0.889E-02**	2.53	0.341E-02	0.59	0.109E-01	1.02	0.747E-03	0.09
INC2	-0.455E-01	-1.33	-0.376E-01	-0.60	0.340E-01	0.31	-0.833E-01	-0.91
INC3	0.534E-01	1.28	0.983E-01	1.21	0.218*	1.64	0.319E-01	0.27
INC4	0.142***	3.01	0.164*	1.84	0.261*	1.77	0.117	0.90
INC5	0.134***	2.66	0.135	1.45	0.112	0.72	0.161	1.18
DSATIS	0.392E-01*	1.92	0.115E-01	0.33	-0.419E-01	-0.66	0.477E-01	0.97
BIGCITY	0.176***	7.53	0.162***	4.15	0.111	1.48	0.192***	3.49
SMALLCITY	0.164E-01	0.63	0.258E-01	0.57	0.510E-01	0.65	0.769E-02	0.12
α	-		1.101***	33.25	-		1.639***	11.27
N	5,011		5,011		5,011		5,011	
Iterations completed	6		16		5		27	
-Log-L	11658.04		9014.738		3088.766		5867.760	
Pseudo R^2	0.164		0.056		0.077		0.046	
Overall % correct	-		-		0.635		-	
OVERDISP;μ[c]	-		7.503***		-		9.898***	
OVERDISP;μ^2[d]	-		7.610***		-		11.182***	
LR test[e]	4569.880***		1069.272***		513.5396		559.636***	
LR test[f]	-		5286.604***		-		3417.240***	
LR test[g]	-		-		-		114.384***	

[a]The estimated parameters (b's) and asterisks indicating significance at 1% level (***), 5% level (**) and 10% level (*).
[b]z is the estimated parameter (b) divided by it's standard error (s.e.).
[c]μ is the t-test of the augmented overdispersion regression $g(\mu_i) = \mu_i$.
[d]μ^2 is the t-test of the augmented overdispersion regression $g(\mu_i) = \mu_i^2$.
[e]Likelihood ratio test of the joint significance of the regressors.
[f]Likelihood ratio test of the Negbin model aginst the Poisson model or of the truncated Negbin model against the truncated Poisson model.
[g]Likelihood ratio test of the Negbin truncated at-zero-Hurdel model aginst the restricted Negbin model.

TABLE 4: ESTIMATION RESULTS: DEPENDENT VARIABLE : FREQUENCY OF CARE WEEKS IN HOSPITAL (COVARIATES INCLUDED).

	RESTRICTED				NEGBIN HURDLE MODELS			
VARIABLE	**POISSON**		**NEGBIN**		**BINARY LOGIT**		**TRUNCATED NEGBIN**	
	b[a]	z=b/s.e[b]	b	z=b/s.e	b	z=b/s.e	b	z=b/s.e
ONE	0.985***	-5.36	-1.247***	-4.93	-2.676***	-8.07	-6.539	-0.02
MALE	-0.214***	-3.90	-0.334**	-2.14	-0.204*	-1.85	-0.212	-0.92
AGE2	0.731***	5.63	0.401	1.58	0.137	0.64	0.358	0.78
AGE3	0.344**	2.46	0.570E-01	0.21	-0.494**	-2.09	0.989**	2.17
AGE4	0.602***	4.18	0.681E-01	0.22	0.102	0.40	0.603	1.17
BADHLTH	2.440***	29.20	2.408***	7.18	1.588***	7.82	1.552***	3.76
FAIRHLTH	1.115***	15.50	1.214***	6.31	0.727***	5.31	0.909***	3.15
LCHRONIC	0.294***	4.29	1.068***	4.23	0.984***	6.11	0.251	0.78
NLCHRONIC	-0.132	-1.24	0.477	1.58	0.675***	3.98	-0.440	-1.08
MARRIED	-0.382***	-4.11	-0.428E-01	-0.14	-0.313*	-1.66	0.514	1.26
FAMSIZE	-0.195***	-5.78	-0.142*	-1.94	-0.653E-01	-1.20	-0.117	-1.14
WORKING	-0.410	-5.68	-0.469***	-2.63	-0.186	-1.39	-0.696E-02	-0.03
EDUCYEAR	-0.127E-01	-1.30	-0.118E-01	-0.58	-0.104E-02	-0.06	0.100E-01	0.28
INC2	-0.250***	-3.05	-0.961E-01	-0.35	0.267	1.37	-1.000**	-2.19
INC3	0.330E-01	0.32	0.101	0.30	0.917***	4.04	-1.446***	-2.73
INC4	0.922E-01	0.75	-0.190	-0.49	0.828***	3.18	-1.825***	-3.02
INC5	0.165	1.22	-0.126	-0.36	0.820***	2.92	-1.838***	-3.18
DSATIS	-0.431	-7.80	-0.707E-01	-0.38	-0.373E-01	-0.34	-0.513*	-1.80
BIGCITY	-0.303***	-4.77	-0.510***	-3.15	-0.263**	-2.01	-0.561**	-2.10
SMALLCITY	-0.439***	-6.06	-0.187	-1.03	-0.176	-1.29	0.268E-01	0.10
α	-		13.340***	16.14	-		5098.5	0.00
N	5,011		5,011		5,011		5,011	
Iterations Completed	9		24		6		97	
-Log-L	4288.504		2108.610		1336.171		712.3706	
Pseudo R^2	0.215		0.066		0.095		0.082	
Overall % correct	-		-		0.913		-	
OVERDISP;μ[c]	-		3.332***		-		14.420***	
OVERDISP;μ^{2}[d]	-		2.537**		-		18.584***	
LR test[e]	2350.060***		298.322***		280.3641***		127.339***	
LR test[f]	-		4359.789***		-		1389.991***	
LR test[g]	-		-		-		126.845***	

[a]The estimated parameters (b's) and asterisks indicating significance at 1% level (***), 5% level (**) and 10% level (*).
[b]z is the estimated parameter (b) divided by it's standard error (s.e.).
[c]μ is the t-test of the augmented overdispersion regression $g(\mu_i) = \mu_i$.
[d]μ^2 is the t-test of the augmented overdispersion regression $g(\mu_i) = \mu_i^2$.
[e]Likelihood ratio test of the joint significance of the regressors.
[f]Likelihood ratio test of the Negbin model aginst the Poisson model or of the truncated Negbin model against the truncated Poisson model.
[g]Likelihood ratio test of the Negbin truncated at-zero-Hurdel model aginst the restricted Negbin model.

TABLE 5: SUB-SAMPLE ESTIMATION RESULTS: DEPENDENT VARIABLE : FREQUENCY OF PHYSICIAN VISITS (COVARIATES INCLUDED).

	GOOD HEALTH				FAIR & BAD HEALTH			
VARIABLE	BINARY LOGIT		TRUNCATED NEGBIN		BINARY LOGIT		TRUNCATED NEGBIN	
	b[a]	z=b/s.e[b]	b	z=b/s.e	b	z=b/s.e	b	z=b/s.e
ONE	0.411**	2.13	0.125	0.61	0.561	1.06	1.231***	5.35
MALE	-0.428***	-6.18	-0.142**	-2.08	-0.443***	-2.60	0.307E-01	0.39
AGE2	-0.205*	-1.75	0.218E-01	0.19	0.156	0.43	0.163	0.91
AGE3	-0.310**	-2.34	-0.195	-1.41	-0.157	-0.43	-0.281	-1.51
AGE4	0.244E-01	0.15	-0.302*	-1.91	0.304	0.75	-0.418**	-2.12
HELBAD	-		-		0.674**	2.49	0.370***	3.85
LCHRONIC	1.360***	4.21	1.326***	5.64	0.851***	3.93	0.500***	5.97
NLCHRONIC	1.425***	8.25	0.810***	6.34	0.495*	1.85	0.182	1.55
MARRIED	0.115	0.95	-0.214	-1.41	-0.376E-01	-0.14	-0.146	-1.17
FAMSIZE	-0.674E-01**	-2.21	-0.443E-01*	-1.68	-0.122	-1.39	-0.542E-01	-1.37
WORKING	0.174**	2.13	-0.558E-01	-0.79	0.171	0.78	0.164E-02	0.02
EDUCYEAR	0.928E-02	0.81	-0.915E-03	-0.08	0.305E-01	0.93	0.14519E-01	1.09
INC2	-0.846E-01	-0.69	-0.456E-01	-0.33	0.523**	2.00	-0.123	-0.94
INC3	0.454E-01	0.31	0.104	0.55	0.966***	2.97	-0.117	-0.76
INC4	0.107	0.67	0.171	0.87	1.011***	2.62	-0.144E-01	-0.08
INC5	-0.334E-01	-0.20	0.126	0.61	0.909**	2.15	0.309	1.58
DSATIS	-0.651E-01	-0.94	0.387E-01	0.54	0.885E-01	0.53	0.106	1.31
BIGCITY	0.114	1.40	0.128	1.63	0.272E-01	0.14	0.283***	3.36
SMALLCITY	-0.136E-01	-0.16	-0.780E-01	-0.85	0.525**	2.26	0.127	1.36
α	-		2.580***	5.82	-	-	1.051***	9.48
N	3,879		3,879		1,132		1,132	
Iterations completed	4		27		5		21	
-Log-L	2583.673		3508.084		489.1011		2326.138	
Pseudo R^2	0.032		0.018		0.065		0.025	
Overall % correct	0.580		-		0.83		-	
OVERDISP;μ[c]	-		6.483***		-		7.355***	
OVERDISP;μ^2[d]	-		6.714***		-		7.723***	
LR test[e]	171.0136***		127.866***		68.256***		121.492***	
LR test[f]	-		1414.481***		-		1887.558***	
LR test[g]	-		-		-		?	

[a]The estimated parameters (b's) and asterisks indicating significance at 1% level (***), 5% level (**) and 10% level (*).
[b]z is the estimated parameter (b) divided by it's standard error (s.e.).
[c]μ is the t-test of the augmented overdispersion regression $g(\mu_i) = \mu_i$.
[d]μ^2 is the t-test of the augmented overdispersion regression $g(\mu_i) = \mu_i^2$.
[e]Likelihood ratio test of the joint significance of the regressors.
[f]Likelihood ratio test of the Negbin model aginst the Poisson model or of the truncated Negbin model against the truncated Poisson model.
[g]Likelihood ratio test of the Negbin truncated at-zero-Hurdel model aginst the restricted Negbin model.

TABLE 6: SUB-SAMPLE ESTIMATION RESULTS: DEPENDENT VARIABLE : FREQUENCY OF HOSPITAL CARE WEEKS (COVARIATES INCLUDED).

	GOOD HEALTH				FAIR & BAD HEALTH			
VARIABLE	BINARY LOGIT		TRUNCATED NEGBIN		BINARY LOGIT		TRUNCATED NEGBIN	
	b[a]	z=b/s.e[b]	b	z=b/s.e	b	z=b/s.e	b	z=b/s.e
ONE	-2.409***	-5.59	-5.517	-0.01	-2.726***	-4.47	-15.069	-0.11
MALE	-0.581***	-3.84	-0.147	-0.33	0.314*	1.86	-0.508E-01	-0.12
AGE2	-0.414E-01	-0.17	0.523E-01	0.08	0.598	1.23	10.767	0.06
AGE3	-0.978***	-3.30	0.556	0.66	0.290	0.59	11.219	0.07
AGE4	-0.191	-0.56	0.804	1.00	0.788	1.55	10.931	0.06
HELBAD	-		-		0.832***	4.32	0.761*	1.83
LCHRONIC	1.550***	4.76	2.239**	2.52	0.776***	4.23	0.327	0.70
NLCHRONIC	1.103***	5.31	-0.745	-1.33	0.159E-03	0.00	-0.443	-0.53
MARRIED	-0.957E-01	-0.36	1.026	1.10	-0.508*	-1.79	0.779E-04	0.00
FAMSIZE	-0.275E-01	-0.42	-0.134	-0.72	-0.158	-1.46	-0.130	-0.39
WORKING	-0.116	-0.69	-0.387	-0.79	-0.152	-0.67	0.132	0.19
EDUCYEAR	-0.243E-01	-0.93	-0.558E-01	-0.95	0.417E-01	1.41	0.881E-01	0.92
INC2	0.251	0.88	-0.929	-1.14	0.281	1.02	-0.985	-1.36
INC3	0.839**	2.57	-1.529	-1.63	0.984***	3.00	-1.353*	-1.69
INC4	0.960***	2.70	-2.805***	-2.67	0.398	0.97	-0.741	-0.64
INC5	0.631*	1.64	-3.053***	-2.92	1.177***	2.77	-0.333	-0.32
DSATIS	0.403E-01	0.27	0.122	0.25	-0.136	-0.82	-0.996*	-1.94
BIGCITY	-0.347	-1.93	-0.638	-1.15	-0.164	-0.83	-0.136	-0.26
SMALLCITY	-0.161	-0.90	-0.539	-0.94	-0.212	-0.98	-0.532	-0.96
α	-		3795.8	0.00	-		886.56	0.010
N	3,879		3,879		1,132		1,132	
Iterations completed	7		82		6		79	
-Log-L	813.9908		246.0504		491.089		438.2075	
Pseudo R^2	0.061		0.139		0.085		0.051	
Overall % correct	0.941		-		0.821		-	
OVERDISP;μ[c]	-		10.582***		-		11.644***	
OVERDISP;μ^2[d]	-		16.699***		-		14.467***	
LR test[e]	106.6446***		79.329***		91.820***		46.710***	
LR test[f]	-		333.4828***		-		857.4218***	
LR test[g]	-		-		-		?	

[a]The estimated parameters (b's) and asterisks indicating significance at 1% level (***), 5% level (**) and 10% level (*).
[b]z is the estimated parameter (b) divided by it's standard error (s.e.).
[c]μ is the t-test of the augmented overdispersion regression $g(\mu_i) = \mu_i$.
[d]μ^2 is the t-test of the augmented overdispersion regression $g(\mu_i) = \mu_i^2$.
[e]Likelihood ratio test of the joint significance of the regressors.
[f]Likelihood ratio test of the Negbin model aginst the Poisson model or of the truncated Negbin model against the truncated Poisson model.
[g]Likelihood ratio test of the Negbin truncated at-zero-Hurdel model aginst the restricted Negbin model.