Health Care Reform and the Number of Doctor Visits - An Econometric Analysis

Summary

Our final project based on the paper which evaluates the German health care reform of 1997. The individual number of doctor visits is the outcome measure. The data set is from German Socio-Economic Panel for the years 1995-1999, 32837 observations are collected.

As the final project for econometrics, we decided to reproduce the work in table 1, 2, 3 and part of table 5 as the pedagogical exercises. Five different models including (1) Poisson regression model, (2) negative binomial regression model, (3) Zero-inflated negative binomial regression model, (4) Hurdle negative binomial model, and (5) Hurdle probit poisson log normal model are fitted to the data set and evaluated by log-likelihood, AIC and BIC. Among all the models, Hurdle negative binomial model demonstrates the best fitness to the data set. Based on the analysis, we found out that XXX. Our conclusions are consistent with the authors of the original paper.

Besides following the analysis in the original paper, we also generate some visualization to describe the data distributions and demonstrate the fitness of the models. The topics we learned in STAT 520 (Ordinary least square, Heteroscedasticity, Two-stage least squares) give us a solid foundation to understand the advanced topics and models such as Generalized linear models, Zero-inflated models, and Hurdle models in the paper.

Introduction

As one of the OECD (The Organization for Economic Co-operation and Development) countries, Germany has a large publicly funded health sector taking a substantial portion of total GDP. A possible explanation for the increase in healthcare spending includes (1) expanding technological and expenditure possibilities in the healthcare services, (2) the surge in demand from the aging population, and incentive structures in the public health sector not promoting the efficient use of resources. One of the significant reforms on the German healthcare system was the reform of 1997, in which the co-payments for prescription drugs were raised by 200%, and the reimbursements of physicians by state insurance were capped. To evaluate the effect of the healthcare reform of 1997, we want to know (1) whether the reform of 1997 is successful and effective and (2) how the effect of reform differs among different groups of the population.

Some background information needs to be elaborated on before we dig into the analysis. The general aim of this reform is to limit healthcare expenditure and its growth rate. The change in co-payments (for prescription) drugs is the most eminent element of the 1997 reform, with measures like an extended exclusion list of drugs not covered at all by social insurance, price ceilings related to the availability of generics, and a binding overall annual budget for drugs and doctor service.

An increased co-payment has a direct fiscal effect of reducing the proportion of cost covered by the insurer. The increased out-of-pocket expenses would disincentive customers from excessive use of prescription drugs and medical resources as a moral hazard.

The analysis of this paper focuses on the disincentivizing effect on prescription drug use by tracking the number of doctor visits by a person during a given time. The rationale for this approach, other than the unavailability of data on drug use, is the correlation between prescription drug use and doctor visit. The implemented policy will increase the expense of drugs, and the patient might try to persuade the doctor to prescribe in larger size or reduce their drug use while not seeing the doctor. Though some other effects might affect the number of visits – one might still see a doctor for diagnosis or advice on non-prescription drugs, the number of such visits would not be affected by the increased co-payment. Thus, an analysis of patient visits could help us get some insights into the effectiveness and success of the 1997 Healthcare reform.

In this report, we will first describe the dataset and how to organize them as the systematic components, then introduce Generalized linear models, Zero-inflated models, and Hurdle models. To this end, five different models are fitted to the data set and we will discuss and interpret the results.

Data

The GSOEP is an ongoing annual household survey and the author selected fives years of data centered around the year of the reform, i.e. 1995-1999. The primary empirical strategy is to pool the data over the five years and estimate the effects of the reforms by comparing the expected number of visits in 1998 and 1996 *ceteris paribus*, i.e., for an individual with given characteristics.

The models that will be estimated in the following sections all include a systematic component (linear predictor) of the type

$$\begin{aligned}x'_{it}\beta =&\beta_0 + \beta_1\text{age}_{it} + \beta_2\text{age}^2_{it} + \beta_3\text{years of education}_{it} + \beta_4\text{married}_{it} + \beta_5\text{household size}_{it}\\&+\beta_6\text{active sport}_{it} + \beta_7\text{good health}_{it} + \beta_8\text{bad health}_{it} + \beta_9\text{self-employed}_{it}\\&+\beta_{10}\text{full-time employed}_{it} + \beta_{11}\text{part-time employed}_{it} + \beta_{12}\text{unemployed}\\&+\beta_{13}\text{equivalent income}_{it} + \beta_{96}(\text{year} = 1996)_{it} + \beta_{97}(\text{year} = 1997)_{it}\\&+\beta_{98}(\text{year} = 1998)_{it} + \beta_{99}(\text{year} = 1999)_{it}\end{aligned}$$

The reference year is 1995. In addition, there are three dummies for the quarter in which the interview took place (winter, autumn, spring).

There are three general channels through which these variables can affect the demand for doctor visits. The first is the underlying health status, the second the budget constraint, and the third the preference formation.

There is no direct health status in the GSOEP. We have the following three sets of proxies instead. Clearly, these are only crude measures of health, and one may want to account for the possibility of additional unobserved heterogeneity to capture any remaining health aspects, as well as other unobserved influences.

- A time-consistent measure of health over 1995–1999 is provided by a subjective self-assessment in response to the question: "How good do you perceive your own health at current?", with responses "very good", "good", "fair", "poor" and "very poor". The two best responses are classified as "good health", the two worst responses as "bad health", with fair health being the reference group.
- Another proxy for health is the age polynomial.
- Finally, engaging in "active sports" (defined as a weekly or higher frequency) acts as a further proxy for good health, although it might have an additional direct effect on the demand for health services as well.

The budget constraint is determined by income and prices. The main price variables are the opportunity costs of a visit to a doctor which, in turn, depend on education level and employment status.

Several of the variables affect more than one aspect at a time. Age, for instance, matters for health, opportunity cost (through the effect of experience on earnings) as well as potentially preferences. Similarly, education is an important factor in determining the optimal investment in health capital (Grossman, 1972). It is not the goal of this paper to disentangle these various transmission channels. Rather, the focus lies on the year dummies, whereas the other right-hand- side variables serve as controls for any effects these variables might have on the changes in visits over time.

Econometric Models for consideration

Since the response variables, the number of the doctor visits, are count data, we first consider generalized linear model (GLM) with Poisson distribution.

Generalized linear model

Generalized linear model (GLM) allows the response variable y to have an error distribution other than a normal distribution. It has three main components: (1) Systematic component (2) Link function (3) Probability distributions. We define $\{(y_i, \mathbf{x}_{i*})\}_{i=1}^n$ as following a generalized linear model based on the exponential family f_θ if

$$y_i \sim f_{\theta_i}, \quad \theta_i = \mathbf{x_{i*}^T}\beta$$

The exponential family is defined as

$$f_\theta(y) = \exp(\theta y - \psi(\theta))h(y)$$

Here θ is called the natural parameter, ψ is called the log-partition function, and h is called the base measure. The distribution with density f_θ is called a one-parameter natural exponential family.

Poisson Regression Model

We take the Poisson distribution as the standard probability distribution for count data

$$P(y_i|\lambda_i) = \frac{\exp(-\lambda_i) \cdot \lambda_i^{y_i}}{y_i!}$$

s.t $\mathrm{E}(y_i \mid \lambda_i) = \mathrm{Var}(y_i \mid \lambda_i) = \lambda_i$

We assume heterogeneous population with covariance x_i in regression, $\lambda_i = \exp(x_i^T\beta)$, $y = (y_1, ...y_N)^T$, $x = (x_1, ...x_N)^T$. The data was randomly sampled by

$$P(y \mid x) = \exp\left[-\sum_{i=1}^{N} \exp(x_i'\beta)\right] \prod_{i=1}^{N} \frac{[\exp(x_i'\beta)]^{y_i}}{y_i!}$$

, and we look for the maximum likelihood (MLE).

The relative change in expected doctor visits as reform effects can be calculated with

$$\begin{aligned}\Delta\%_{(98,96)} =& 100 \times [\frac{\mathrm{E}(y_{i,98}|x)}{\mathrm{E}(y_{i,96}|x)} - 1] \\ =& 100 \times [\exp(\beta_{98} - \beta_{96}) - 1]\end{aligned}$$

It should be noted that the Poisson Model has a few shortcomings:

1. The assumption exclude the case of unobserved heterogeneity
2. It ignore the Panel Structure of the data and the Standard Errors need to incorporate possible serial correlation (unless we can assume constant individual effect over time)
3. The single index structure of the Poisson Regression model implies that the distribution is determined w.r.t the given mean, and we cannot evaluate the effect of the reform over the distribution (i.e cannot how it affect different sections of the population)

The random effect and fixed effect models In statistics, a random effects model, also called a variance components model, is a statistical model where the model parameters are random variables. It assumes that the data from a hierarchy of different populations whose differences relate to that hierarchy and the hetrogeneity is uncorrelated with the independent variables. Random effect models assist in controlling for unobserved heterogeneity when the heterogeneity is constant over time and not correlated with independent variables. This constant can be removed from longitudinal data through differencing, since taking a first difference will remove any time invariant components of the model. [[From wikipedia https://en.wikipedia.org/wiki/Random_effects_model]]

On the other side, a fixed effects model is a statistical model in which the model parameters are fixed or non-random quantities. It assumes that the individual-specific effects are correlated with the independent variables. It assists in controlling for omitted variable bias due to unobserved heterogeneity when this heterogeneity is constant over time. This heterogeneity can be removed from the data through differencing, for example by subtracting the group-level average over time, or by taking a first difference which will remove any time invariant components of the model. [[From wikipedia https://en.wikipedia.org/wiki/Fixed_effects_model]]

Negative binomial regression Model

To model the count data, we can also use Negative binomial regression. It is a generalization of Poisson regression. Negative binomial distribution can be considered as a generalization of Poisson distribution by including a gamma noise variable which has a mean of 1 and a scale parameter of ν. The poisson-gamma mixture distribution that results is

$$\Pr(Y = y_i|\mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1)\Gamma(\alpha^{-1})}(\frac{\alpha^{-1}}{\alpha^{-1}})$$

, where $\mu_i = t_i\mu$ and $\alpha = \frac{1}{\nu}$.

Zero-inflated negative binomial model

Zero-inflated models have been developed to cope with zero-inflated outcome data with over-dispersion (negative binomial) or without (Poisson distribution). In this project, zero-inflated negative binomial regression is applied on the dataset.

A zero-inflated (ZI) model assumes that the zero observations have two different origins: "structural" and "sampling". The sampling zeros are due to the usual negative binomial distribution, which assumes that those zero observations happened by chance. Zero-inflated models assume that some zeros are observed due to some specific structure in the data. [[https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3238139/]]

For example, in the problem of the original paper, the number of the doctor visits is the outcome, some participants may visit zero in certain year because they were cured; these are the structural zeros since they cannot exhibit the reform of the healthcare. Others participants have zero number of visit because they purchased more medicines because of the healthcare reform and it changed their behavior. To this end, the number of doctor visit is assumed to be on a negative binomial distribution that includes both zero (the "sampling zeros") and non-zero counts.

Hurdle Models

In contrast to ZI models, a hurdle model assumes that all zero data are from one "structural" source. The positive (i.e., non-zero) data have "sampling" origin, following either truncated Poisson or truncated negative-binomial distribution[[https://www.sciencedirect.com/science/article/abs/pii/0304407686900023]]. The hurdle model has been popular in the health literature, in part because it can be given a structural interpretation that seems to agree well with the intuition of a dual decision structure of the demand process.

For example, consider a study of of the original paper. It is also reasonable to assume that it is most likely that healthy people will zero number of doctor visit and people who feel uncomfortable or have not finish the treatment will have non-zero number of doctor visits in a year. Hence the zero observations can come from only one "structural" source, the patients. If a subject is considered a patient., they do not have the "ability" to score zero number of doctor visits in a whole year and will always have a positive number in a hurdle model with either truncated Poisson or truncated negative binomial distributions.

Define $d_i = 1$ if a person does not see a doctor in the period and $d_i = 0$ if he/she sees a doctor, s.t. $d_i = 1 - \min(1, y)$, the probability function of the hurdle model is

$$f(y_i) = f_{1i}^{d_i} \left[(1 - f_{1i}) f_T(y_i \mid y_i > 0)\right]^{1-d_i}$$

, with $f_{1i} = P(d_i = 1)$, $f_T(y_i|y_i > 0) = \frac{f_2(y_i)}{1-f_{2i}(0)}$.

For f_1, we have the following hurdle functions as candidates

1. Poisson: $\exp(-\exp(x_i^T\gamma))$
2. Negative Binomial, Type 1: $(1+\theta)^{-\exp(x_i^T\theta)/\theta}$
3. Negative Binomial, Type 2: $[\alpha/(\exp(x_i^T\alpha) + \alpha)^{\alpha}]$
4. Logit: $\frac{\exp(x_i^T\gamma)}{1+\exp(x_i^T\gamma)}$
5. Probit: $\Theta(x_i^T\gamma)$

For f_2, we have

1. Poisson
2. Negative Binomial
3. Poisson-log-normal $p(x) = (\exp(x\mu + x^2\sigma/2) \cdot (2\pi\sigma)^{-1/2})/x! \cdot g(y)$

Generalized linear model

Results

Table 1

We first reproduce table 1 to generate the summary statistics for the variables of the analysis. One can have the following observations

- The average number of doctor visits per quarter declined from 2.66 to 2.35 between 1996 and 1998.
- The decline of the number of doctor visit between 1995 and 1996 is 1%. The increase between 1998 and 1999 is 2%. Therefore, the large drop (7.8%) between 1997 and 1998 in the number of visits clearly coincides with the timing of the reform.
- The increase of the number of doctor visit between 1998 and 1999 also supports the behavior hypothesis. It is because that the 1997 reform enjoyed only a short lifespan. The partial repeal of the 1997 reform was one of the first items on the political agenda of the new government selected in 1998, and a new law lowered the co-payments by between DM 1 and DM 3, effective January 1, 1999.
- The proportion of people reporting good health from 56 to 60% while the proportion reporting bad health decrease from 14 to 13%. This can be the partial reason supporting the idea that the 1997 reform did not worsen the general health status.

	1995	1996	1997	1998	1999
Total Observation	6790.000	6555.000	6480.000	6781.000	6231.000
No. doctor visits	2.687	2.657	2.553	2.353	2.391
Relative change in %	NA	-0.011	-0.039	-0.078	0.016
No. doctor visits (0/1)	0.348	0.328	0.352	0.372	0.346
Age	38.081	38.203	38.471	38.731	38.919
Unemployed (0/1)	0.085	0.084	0.092	0.085	0.075
Active sport (0/1)	0.295	0.247	0.262	0.307	0.266
Good health (0/1)	0.568	0.562	0.581	0.595	0.580
Bad health (0/1)	0.145	0.138	0.134	0.127	0.129

Table 2

To fit and describe the number of doctor visit as the count data, we first applied Poisson regression while considering the random and fixed effects. The log-likelihood of the models are as follows.

- The pooled model: -86566.18.
- The random effect model: -70176.72.
- The fixed effect model: -42738.86.

The estimations of coefficients and stander errors for the basic Poisson model, with and without individual specific effects, are displayed in Table II.

	Pooled Estimate	Pooled Robust SE	Random Effects Estimate	Random Effects Robust SE	Fixed Effects Estimate	Fixed Effects Robust SE
Age	-0.01057	0.00660	-0.00927	0.00575		
Age^2	0.00016	0.00008	0.00019	0.00007		
Male	-0.20852	0.02091	-0.29480	0.02028		
Education	-0.00576	0.00370	-0.00424	0.00408		
Married	0.08078	0.02226	0.07733	0.01809	0.10358	0.02507
Househould size	-0.05221	0.00777	-0.05547	0.00596	-0.05991	0.00850
Active Sport	0.04657	0.01935	0.00614	0.01257	-0.01015	0.01417
Good health	-0.61088	0.01901	-0.50135	0.01183	-0.43002	0.01277
Bad health	0.81311	0.02250	0.62252	0.01216	0.56610	0.01288
Social assistance	0.08610	0.04444	0.04053	0.02640	0.01763	0.02891
Log(income)	0.09312	0.02286	0.01737	0.01598	-0.01659	0.01924
Year=1996	0.00122	0.02676	-0.01419	0.01110	-0.01022	0.01122
Year=1997	-0.03024	0.02689	-0.05027	0.01143	-0.04030	0.01153
Year=1998	-0.10481	0.02681	-0.10615	0.01187	-0.08439	0.01198
Year=1999	-0.09868	0.02726	-0.10685	0.01233	-0.07929	0.01246

One can make the following observations from Table II

- Men have less doctor visits than women.
- The expected number of doctor visits is u-shaped in age because `age` is negative.
- The health indicators (`Good health` and `Bad health`) have the largest effect among all variables.
- There is a statistically significant decline in the expected number of doctor visits between 1996 and 1998 in all three model specifications.

Table 3

Four additional models were applied on the same data set. Likelihood ratio tests clearly reject the Poisson model against the alternative models with unobserved heterogeneity. AIC is used to pick the best model among all five models. We also calculated BIC as the alternative criteria and it supported the same conclusion as AIC. Zero-inflated negative binomial and hurdle negative binomial models are the best two models.

Models	Log.likelihood	AIC	BIC
Poisson	-86566.18	173176.4	173361.1
Negative Binomial	-64611.55	129269.1	129462.3
Zero-inflated Negative Binomial	-64369.56	128829.1	129207.1
Hurdel Negative Binomial Model	-64254.90	128599.8	128977.8
Hurdle Probit Poisson Log Normal Model	-77976.35	156040.7	156410.3

Table 5

The relative changes in expected doctor visits from 1996 to 1998 for each models are as follows. The basic models (Poisson and Negative Binomial models) estimate the reform effect as -10% and -8%. However the structural models (Zero-inflated negative binomial and Hurdle negative binomial models) estimate the reform effect as -6% and -3%. The differences between the two kinds of models support the idea that the survey, which was conducted in the hospital, was over-represented by patients with more doctor visits.

Models	Delta % (96, 98)
Poisson	-10.060%
Negative Binomial	-8.890%
Zero-inflated Negative Binomial	-6.321%
Hurdel Negative Binomial Model	-3.482%

Discussion

As an expansion of the author's work, we tried to visualize the distribution of the number of the visits of doctors per year with a box plot and violin plot (Figure 1). But the outliers make it difficult to tell the differences of the median of numbers of visits among the years. The density histogram (Figure 2) of the numbers of doctor visits is also generated. We can observe that year 1998 has a slightly more zeros than other years. But any of the above visualizations are significant enough to be used as a evidence showing the effect of 1997 reform, which makes the analysis conducted in the original paper even more important. This visualization exercise help us realize that for the panel data with short time series and many observations. Regression is a more powerful tool to intercept the data than plots.

As another expansion and alternative to the Hurdle models in the original paper, we applied a zero-inflated negative binomial model (ZINB) on the data. Based on Table 3, ZINB outperformed basic models (POisson and Negative binomial models) but fitted slightly worse than Hurdle Negative binomial model. It supports the idea that Hurdle assumption aligns better to the data. In other words, we tends to believe that all zero data are from one "structural" source.

The index models used in this research can be applied to study the effect of other medical healthcare events. From 1990 - 2015, multiple attempts were made to improve the solidarity-based system through competition to further improve economic efficiency of the healthcare system and promote technological advancement. (https://doi.org/10.1016/S0140-6736(17)31280-1) If policy objective is on the effect of promoting efficient and responsible use of healthcare services, we might apply the same analysis to the Act to Strengthen Competition in Statutory Health Insurance in 2007 and the Statutory Health Insurance Care Structures Act, given that the primary policies of 2007 reform are 1) mandatory universal coverage, 2) introduction of a uniform contribution

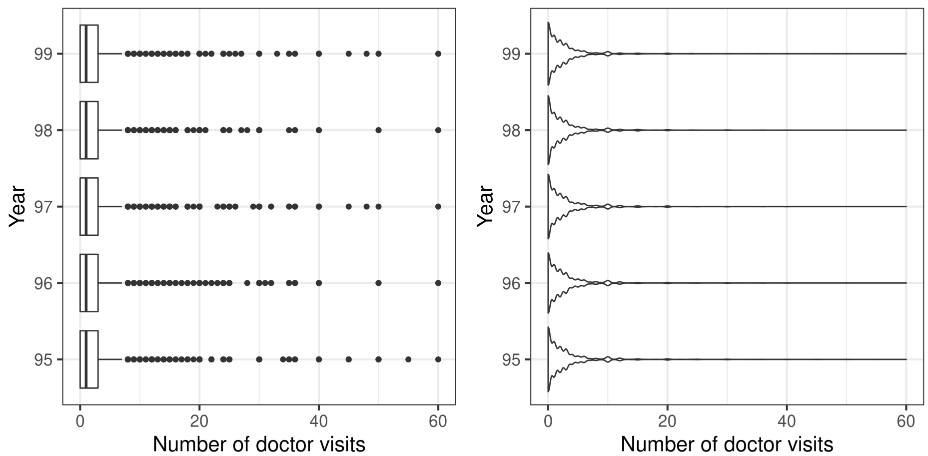

Figure 1: Box plot and violin plot of the numbers of doctor visits per year

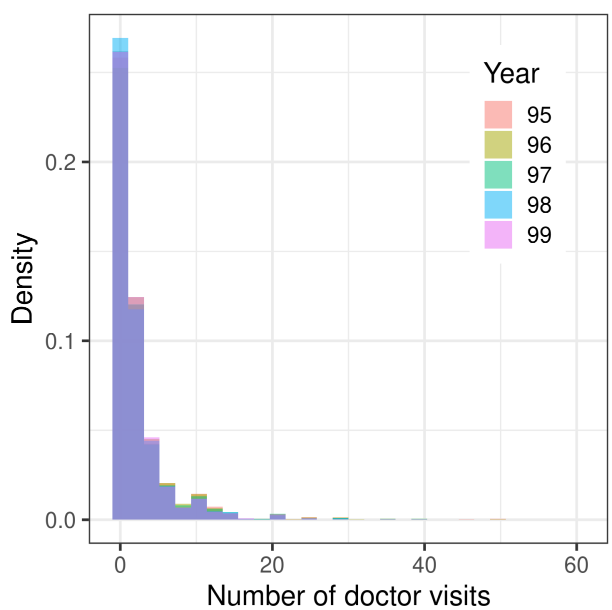

Figure 2: Density Histogram of the numbers of doctor visits per year

rate, a central reallocation pool and resource allocation to sickness funds according to a morbidity-based risk structure compensation scheme, 3) choice of tariffs in statutory health insurance, while the 2011 Acts are more focused on supplementary premiums and reimbursement for new pharmaceutical products. It should be noted that to evaluate the effect of both reforms, in addition to variables we used in previous analysis, we might want to incorporate more economic variables related to economic growth, business cycles, considering the 2008 recession and 2010s European Sovereign Debt Crisis.

Appendix

The calculation of table 1

The calculation of table 2

The calculation of table 3

The calculation of table 5

Box, violin, and histogram plots