On the Anatomy of Adverse Selection in Health Insurance Market: Evidence from the MEPS

Cinzia Di Novi

AIES
Associazione Italiana di Economia Sanitaria

"Of course the policy is automatically cancelled should either of you need medical treatment."

Rational

[…Since uncertainty, much of an individual's demand for health care is not steady, but irregular and unpredictable. This implies that the costs of health care act as a random deduction from an individual's income. Therefore, under uncertainty, risk-averse individuals demand risk-bearing goods, such as health insurance, to safeguard their income against possible shocks…]

(Cagatay, 2004, " The Effects of Uncertainty on the Demand for Health Insurance", *Journal of Risk and Insurance*, 71: 45-61.).

Rational

- It is well known, however, that health insurance coverage reducing an individual's marginal cost of medical care inputs, leads to use additional medical services: an insured individual, in fact, may consume more medical services and have a greater expenditure compared to an uninsured case (i.e. *moral hazard effect*) (Leibowitz, 2004).

- Insurance choice itself may be affected by planned medical expenditure and expectations about medical care utilization (i.e. *adverse selection effect*).

Rational

- This paper focus in particular on this last effect.
- In health insurance market ***adverse selection*** may occur when consumers' true health-cost risk is private information: insurance company may know that consumers vary in the level of risk, but, on principle, is not able to discern who are high and who are low risk profile individuals within a group of potential insured. (Akerlof, 1970; Rothschild and Stiglitz, 1976).

Rational

- Identifying risks accurately is not an easy task and requires that insurance company incurs some costs. Insured parties are heterogeneous in terms of expected costs and have more information about their risks.

- Naturally, high-risk individuals are not encouraged to "reveal" their risk to the insurance company; this asymmetry is a serious problem since may lead insurance company to face large differences in expected health costs due to heterogeneity in demographics and the incidence of illness.

Objective

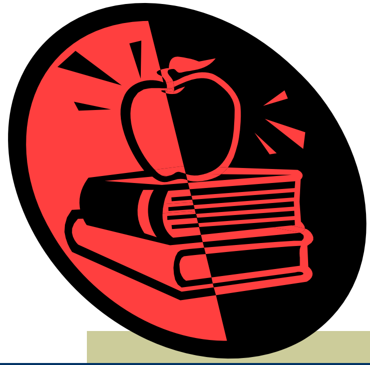
- In spite of the extensive theoretical interest on the adverse selection, there is little empirical evidence on the extent of the problem.
- The goal of this paper is to test empirically for adverse selection in the U.S. health insurance market.

Hypothesis

- Risky individuals who expect high health care costs would tend to purchase insurance with higher premium but lower excesses since they are more likely to be claiming on a regular basis.

- On the other hand lower risk users, who expect low costs, would prefer a less complete insurance, with a lower premium and a higher excess in the unlikely event that they have to claim.

Test



- The "**positive correlation property**" between the individual riskiness and the completeness of a health insurance plan forms the basis for our empirical test for adverse selection.

Data

- The test is based on the 2003/2004 Medical Expenditure Panel Survey - Household Component (MEPS-HC) data used in conjunction with the previous year's National Health Interview Survey (NHIS) data.

- MEPS data contain detailed information on health care consumption and demographics including age, sex, marital status , income, work status and geographic location. In addition the data contain information on the respondents' health status, health charges and payments, access to care, health conditions, health insurance coverage.

Data

- NHIS data provides rather detailed information about health status, diseases, life-style, education, and other individual characteristics.

- After correcting for the missing values, the sample was reduced to 890 individuals resulting in 1780 observations.

- Observations containing veterans and individuals who are covered by Champus/ ChampVa insurance are removed from the data set since their medical services demand and access to medical services distinctly differs from the general population.

Risk-Profile Variables

- To test if high risk individuals buy more generous plan and whether a separating or a pooling equilibrium best characterized the market, we classify the individuals on the basis of their high and low risk profile.
- Individuals are classified as being low risk

If their perceived health status is: excellent, very good, good.

If they do not suffer from hypertension (high blood pressure)

If they have a healthy life-style: they do not smoke, do not usually consume heavy drinks, practice vigorous physical activity more than once per week and if their reported BMI is less than 25.0000.

Methodology

- We measure health insurance plan completeness by using health insurance reimbursement that is the vertical difference between total health expenditure and out-of-pocket expenditure on health care paid by consumers.

- Health insurance reimbursement, however, is only defined for a subset of individuals from the overall population since we observe it only for those who participate in insurance and have positive health care expenditure.

Methodology

- Thus, the model may suffer from sample selection bias and straightforward regression analysis may lead to inconsistent parameters estimate.
- Another problem that arises from the estimation is the presence of unobserved heterogeneity in the equations of interest.

Methodology

- Wooldridge (1995) has proposed an estimator which deals with both sources of estimation bias.
- We extend this estimation method to the case in which selectivity is due to two sources rather than one (participation in insurance and participation in health care expenditure).
- The nature of the test is similar to the one in Browne and Doerpinghaus (1993).

Wooldridge's(1995) two-step estimation

Concerning the health insurance reimbursement model, we consider the following characterization of the Wooldridge's sample selection model where the selectivity bias is a function of two indices

$$\left.\begin{aligned} d^{*}_{it_1} &= z_{it_1}\gamma_1 + \mu_{i_1} + u_{it_1} \\ d_{it_1} &= 0 \quad if \quad d^{*}_{it_1} \le 0 \\ d_{it_1} &= 1 \quad if \quad d^{*}_{it_1} > 0 \end{aligned}\right\}$$

Selection equation(1) .Let d_{it_1} be an unobserved variable denoting the insurance participation decision.

Wooldridge's(1995) two-step estimation

$$d^*_{it_2} = z_{it_2}\gamma + \mu_{i_2} + u_{it_2}$$
$$d_{it_2} = 0 \quad if \quad d^*_{it_2} \leq 0$$
$$d_{it_2} = 1 \quad if \quad d^*_{it_2} > 0$$

Selection equation(2) . Let d_{it_2} be an unobserved variable denoting the health care expenditure participation. decision.

$$y^*_{it} = x_{it}\beta + \alpha_i + \varepsilon_{it}$$
$$y_{it} = y^*_{it} \quad if \quad d_{it} = 1$$
$$y_{it} \quad not \quad observed \quad otherwise$$

Outcome/Primary equation. Let y_{it} be an unobserved variable denoting the natural logarithm of health insurance reimbursement.

Wooldridge’s(1995) two-step estimation

- The sample selection is now based on two criteria.
- The method of estimation relies crucially on the relationship between u_{it_1} and u_{it_2} .
- In particular, the estimation depends on whether the two error terms are independent or correlated, that is whether or not $Cov(u_{it_1}, u_{it_2}) = 0$.
- The simplest case is when the disturbances are uncorrelated (Maddala,1983, Vella, 1998).
- In that cases we can easily extend the Wooldridge’s two-step estimation method to this model.

Wooldridge's(1995) two-step estimation

The correction term to include as regressor in the primary equation is:

$$E\left[v_{it} \middle| z_{it}, d_{it_1} = 1, d_{it_2} = 1\right]$$
$$= \rho_{t_t} \lambda_1\left(z_{i1_1}\gamma_{1_1} + \ldots + z_{it_1}\gamma_{t_1}\right) + \rho_{t_2}\lambda_2\left(z_{i1_2}\gamma_{1_2} + \ldots + z_{it_2}\gamma_{t_2}\right)$$

We estimate the following model

$$y_{it} = x_{i1}\psi_1 + \ldots + x_{it}\psi_t + x_{it}\beta +$$
$$+ (\phi_{t_1} + \rho_{t_1})\hat{\lambda}_1(\bullet) + \left(\phi_{t_2} + \rho_{t_2}\right)\hat{\lambda}_2(\bullet) + e_{it}$$

Wooldridge's(1995) two-step estimation

- The procedure consists in first estimating, for each period, by two single cross-sectional probit models, the selection equation one and the selection equation two.

- Then, the two corresponding Inverse Mills Ratio can be imputed and included as correction terms in the primary equation.

- Thus, by fixed effect or pooled OLS, estimate of the resulting primary equation corrected for selection bias can be done for the sample for which $d_{itj} = 1$.

Wooldridge's(1995) two-step estimation

In the case u_{it_1} and u_{it_2} are correlated, so that , $Cov\left(u_{it_1}, u_{it_2}\right) = \sigma_{12}$
"… the expression get very messy…" (Maddala, 1983) and we have to use for each period cross-sectional bivariate probit methods to estimate γ_{it_1} and γ_{it_2}. Further,

$$E\left[v_{it} \middle| z_{it}, d_{it_1} = 1, d_{it_2} = 1\right] = \rho_{t_1} M_{12} + \rho_{t_2} M_{21}$$

where:

$$M_{ij} = \left(1 - \sigma_{12}\right)^{-1}\left(P_i - \sigma_{12} P_j\right)$$

$$P_j = \frac{\int_{-\infty}^{z_{it1}\gamma_{t1}} \int_{-\infty}^{z_{it2}\gamma_{t2}} u_{it_j} f\left(u_{it_1}, u_{it_2}\right) du_{it_1} du_{it_2}}{F\left(z_{it_1}\gamma_{t_1}, z_{it_2}\gamma_{t_2}\right)}$$

Before starting the estimation we run a preliminary cross-sectional bivariate probit…

Bivariate Probit Models for Health Expenditure and Insurance Participation

- The null hypothesis of $Cov\left(u_{it_1}, u_{it_2}\right) = 0$ is not rejected. The table below shows the correlation coefficients and the p-value for each year:

Dependent Variables	*pho*	*p-value*
Positive Expenditure/ Be Insured 2003	-0.5299	0.260
Positive Expenditure/ Be Insured 2004	-0.9496	0.541

- Since the error terms are independent we can deal with the above model as independent equations (Maddala, 1983):

Conclusions

We find no systematic relation between illness of individuals and insurance choice.

We think that a possible explanation can be found in the so called "cream skimming" practise: health plans may have an incentive to alter their policy to attract the healthy and repeal the sick (Newhouse, 1996; Ellis, 1997).

Then, insurers may practice a kind of "reverse adverse selection": they would try attempt to increase their profits by refusing to write policies for the worst risks in an insurance pool (see Siegelman, 2004).

Conclusions

These strategic behavior can take a variety of forms including: designing insurance benefits packages in such a way as to be more attractive to healthy persons than unhealthy one for instance by excluding particular prescription drugs, offering numerous pediatrician (families with children are better risks) or by excluding cancer specialist visits.

If health plans cream healthy individuals, those who are enrolled in health insurance are relatively healthy people and this lead to the failure of the correlation test.