

## THE IMPACT OF SUBSIDIZED HEALTH INSURANCE ON THE POOR IN COLOMBIA: EVALUATING THE CASE OF MEDELLÍN

ANDRÉS RAMÍREZ HASSAN \*  
JOHNATAN CARDONA JIMÉNEZ †  
RAMIRO CADAVID MONTOYA ‡

### Abstract

This paper uses count and binary data models with an endogenous dummy variable to evaluate the effect of the subsidized health care program in Medellin (Colombia). The subsidized program, which primarily covers poor people, is found to have a significant impact on the use of preventive medical care and hospitalization that might have a negative impact on the financial statements of the program. Specifically, econometric estimations of health care utilization indicate that there is both selection and moral hazard. These facts imply that the program can improve its coverage if mechanisms are created to lower the individual moral hazard effect.

**Keywords:** Binary Models; Count Models; Endogenous Dummy Regression; Program Evaluation.

### Resumo

Este trabalho utiliza modelos de dados binários e de contagem com uma variável indicadora endógena para avaliar o efeito do programa de saúde em Medellín (Colômbia). Encontrou-se um impacto significativo do programa subsidiado, que abrange principalmente a população pobre, sobre o uso de assistência médica preventiva e de internação, o que pode ter um impacto negativo sobre as situações financeiras do programa. Especificamente, estimativas econométricas da utilização de cuidados de saúde indicam que há seleção e risco moral. Esses fatos sugerem que o programa pode melhorar sua cobertura se forem criados mecanismos para diminuir o efeito de risco moral individual.

**Palavras-chave:** Avaliação de programas; Modelos binários; Modelos de contagem; Regressão com dummy endógena.

**JEL classification:** I18, I38, C21

---

\* Universidad EAFIT. E-mail: [aramir21@eafit.edu.co](mailto:aramir21@eafit.edu.co)

† Universidad Nacional de Colombia. E-mail: [jcardonj@unal.edu.co](mailto:jcardonj@unal.edu.co)

‡ Universidad EAFIT. E-mail: [ramirocadaavid@eafit.edu.co](mailto:ramirocadaavid@eafit.edu.co)

## 1 Introduction

Health care utilization and health insurance are intimately related because of moral hazard and self selection. The demand for health care is conditioned by the health insurance status of the user. However, the insurance decision itself depends upon the expected future use of health services Cameron et al. (1988). Specifically, the subsidized health insurance program reduces the effective price of medical care; this may lead to “overutilization” by the insured users. Additionally, the future use of health services may induce people to enroll in the subsidized health insurance program. This phenomenon is known as simultaneous equation bias or endogeneity bias. The problem is that the error term of the medical care utilization is correlated with a dummy variable that determines the participation in the program. In order to take into account simultaneity, we use count and binary models based on Full Information Maximum Likelihood that accommodate endogeneity due to self selection associated with treatment effects (Amemiya 1978, Terza 1998).

Recently, a great effort has been made to provide basic health care insurance to the poor in developing countries. In Colombia, the Philippines and Vietnam, for example, the poor are enrolled in a national social health insurance scheme, which is financed by taxpayers. In China and Mexico, by contrast, households not covered by formal sector programs have the option of enrolling in a separate subsidized public health insurance program Wagstaff et al. (2007). However, the potential gains, because of distributional effects, can be lost due to efficiency losses caused by distortion in users’ behavior associated with moral hazard and self selection. As a result, policy makers have to evaluate the programs in order to identify problems with their implementation.

Evaluations of subsidized health insurance programs in developing countries are generally based on propensity score matching Saadah et al. (2001), Trujillo et al. (2005), Urdinola & Jain (2006), Wagstaff (2007), Wagstaff et al. (2007), Pita et al. (2008). The empirical evidence based on propensity score matching indicates that subsidized health insurance programs have resulted in a net increase in utilization for the poor beneficiaries. Nevertheless, the level of health care utilization remains low in developing countries. However, matching estimators of treatment effects are useful when the process of selection in treatment is based on observable variables, i.e., the key assumption is that unobservable variables play no role in the treatment assignment and outcomes determination. This is known as conditional independence assumption Cameron & Trivedi (2005). If the conditional independence assumption is not met, the estimated treatment effect is biased and inconsistent. In recent years, instrumental variable methods have been advocated in order to account for unobserved confounding variables, i.e., endogeneity problems. Specifically for the Colombian case, Trujillo et al. (2005) and Gaviria et al. (2006) used instrumental variable methods; they found that the subsidized insurance program greatly increased medical care utilization among the country’s poor. However,

“...Despite the substantial increase in public expenditure on health care and the increase in the proportion of population with health insurance, many problems persist. On the one hand, the implantation of a scheme of subsidies to demand has not been ac-

accompanied by a dismantling of subsidies to supply, which has led to a doubling in expenditure and a multiplication of inefficiencies.” Gaviria et al. (2006, p. 4)

These structural problems recently caused the Colombian government to declare a social emergency because of the health care system going bankrupt.

This paper evaluates the impact of Colombia’s subsidized health insurance program on preventive health care utilization and hospitalization services of Medellín’s citizens, and it attempts to contribute to the more general literature on the impact of subsidized health insurance programs implanted in developing countries. In particular, we use Poisson and Binary models that take into account endogeneity due to unobserved characteristics. Our econometric estimations of health care utilization indicate that there is both selection and moral hazard. These phenomena invalidate the evaluation of subsidized health programs based on propensity score matching, which is the most popular technique used in these circumstances. Specifically, it is found that people enrolled in this subsidized program visit less frequently a doctor compared with uncovered individuals. On the other hand, covered people have a higher probability of hospitalization than individuals who are not in this program. This result has bad consequences for financial statements of the subsidized health care program.

The paper is organized as follows. The first section provides a review of the Colombian health care systems. The methodology used to evaluate the impact of the subsidized health insurance on the poor in Colombia is sketched out in section two. Section three presents the evaluation results of the subsidized health care program in Medellín city, and finally, we show some concluding remarks.

## 2 Colombian health care systems

Up until the 80s, the Colombian health care system was based on three separate schemes: the private scheme, which provided health services to the wealthiest segment of the population through private health insurance plans; the public scheme, which had the task of providing health services to the poorest part of the population who were not protected by medical insurance; and the social security scheme, which aimed to provide basic health services to two different groups of the population, the formal workers belonging to the private sector through the Social Security Institute (Instituto de Seguridad Social) and the public sector workers through the Social Benefit Societies (Cajas de Previsión Social). This health care system had its boom between 1975 and 1984 but fell into decay after that due to the fiscal crisis during 1984 that reduced the contributions assigned to this sector. Despite the aforementioned boom, the system never achieved near acceptable goals, mainly because inequalities persisted (both between regions and strata), low levels of insurance coverage and a high inefficiency in the public provision of health services Gaviria et al. (2006).

During the first half of the 1990s, the Colombian government carried out profound reforms in health care system that reduced its participation in the industry. This allowed for regulated competition among private firms, separated the provision of health insurance from supplying health services, and decentralized subsidized programs for the local governments. The schemes

that arose out of these reforms are a contributory social insurance program financed by mandatory taxes, and a subsidized program aimed at covering the lowest income and the most vulnerable of the population, financed by both local and central government resources. These reforms were based on five principles, according to the 1993 Law 100: efficiency, universality, solidarity, integrity and participation, which tried to achieve four objectives: increase coverage, increase solidarity through cross subsidies from the contributive to the subsidized program, improve efficiency by allowing competition in the industry and changing the supply side subsidies to demand side subsidies Congress (1993).

The new health system required people affiliated to the contributive program to pay 12% of their income (two thirds paid by the employer and one third by the employee), collected by an insurer chosen by the employee. The subsidized program is financed by a constant percentage of the aforementioned contribution, transfers from the central government and resources owned by the local governments and regional entities. The municipalities are responsible for identifying the beneficiaries of the subsidized program, who pay a coinsurance rate in order to rationalize the use of medical services. Additionally, there is still a supply side subsidy paid to ESS (Empresas Solidarias de Salud), which are nonprofit medical care providers.

Compared to the reform's explicit objectives that led to the actual health care system, this has not been successful. It has targeting problems, due to a significant number of low income households which do not receive the benefits from the subsidized program when some high income families do. This happens because of imperfect monitoring, corruption and limitations in administrative capabilities to identify plausible beneficiaries. It also has coverage problems because of the slow growth of people in the contributive program, transfers from the central government and contributions from the regional entities Gaviria et al. (2006). Additionally, the transition of supply side to demand side subsidies is still incomplete, maintaining inefficiencies due to public hospitals which are not capable of selling services, thus displaying a structural deficit in their budget. Along with this, the introduction of the subsidized program has led to a growth in the number of public hospitals which lower the levels of occupation, resulting in an underuse of these facilities. The biggest problem of the failure to make the transition between supply side to demand side subsidies is that the first extracts resources from the latter.

### **3 Methodology**

The basis for evaluating programs is the counterfactual analysis where an outcome is observed for all individuals but in different states. Specifically, we consider a model in which health care utilization is observed for all individuals, but in only one of the two possible states; covered or uncovered by the subsidized program. Individuals self select into a given program due to expected future use of preventive care and hospitalization might affect the decision to enroll in the scheme, this phenomenon has the potential for causing endogeneity. Given that our measures of health care utilization are number of preventive visits and hospitalization, we must specify statistical models that have the same support as our dependent variables, specifically count and binary responses, respectively. Therefore, we use Poisson and Binary models

where self selection due to unobserved characteristics is being additionally taken into consideration. Specifically, we use the methodology developed by Terza (1998) and Miranda (2004) in the case of number of preventive doctor visits, and we apply the procedures proposed by Amemiya (1978), Green (2003) and Miranda & Rabe-Hesketch (2006) which account for endogeneity in binary response models, which is the case in hospitalization.

Regarding the number of preventive visits, we follow the discussion of Terza (1998) and Miranda (2004) where the  $i$ th individual is considered from a random sample  $I = \{1, 2, \dots, N\}$  whose dependent variable is the number of preventive visits to the doctor in the last year ( $y_i = 0, 1, \dots$ ). The conditional probability density function of the count dependent variable is Poisson whose mean equals to  $Exp\{x'_i\beta + d_i\gamma + e_i\}$ . This one depends on a vector of explanatory variables  $x_i$ , a binary switching variable ( $d_i = 0, 1$ ) which indicates the state of  $i$ th individual, covered or uncovered by the subsidized health insurance program, and a random component  $e_i$  that accounts for unobserved heterogeneity. Given a vector of explanatory variables  $z_i$  (which may contain some or all the elements of  $x_i$ ),  $d_i$  is characterized by an index process

$$d_i = \begin{cases} 1, & z'_i\alpha + v_i > 0 \\ 0, & z'_i\alpha + v_i \leq 0 \end{cases} \tag{1}$$

Suppose that  $e_i$  and  $v_i$  are jointly Normal,

$$\begin{bmatrix} e_i \\ v_i \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \sigma\rho \\ \sigma\rho & 1 \end{bmatrix} \right) \tag{2}$$

where  $\rho$  accounts for correlation between the response variable and the treatment indicator.

After some algebra and given a change of variable  $\eta_i = e_i/\sigma\sqrt{2}$ , the joint conditional probability density function of  $y_i$  and  $d_i$ , given  $x_i$  and  $z_i$ , may be expressed as

$$f(y_i, d_i | z_i, x_i) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} f(y_i | z_i, x_i, d_i, \sqrt{2}\sigma\eta_i) \begin{bmatrix} d_i\Phi_i^*(\sqrt{2}\sigma\eta_i) + \\ (1-d_i)(1-\Phi_i^*(\sqrt{2}\sigma\eta_i)) \end{bmatrix} Exp\{-\eta_i^2\} d\eta_i \tag{3}$$

where  $f(y_i | z_i, x_i, d_i, \sqrt{2}\sigma\eta_i)$  is the conditional probability density function of  $y_i$  given  $z_i, x_i, d_i$  and  $e_i$ ;  $\Phi(\cdot)$  is the cumulative distribution function of a Normal standard variable; and  $\Phi_i^*(\sqrt{2}\sigma\eta_i) = \Phi\left(\frac{z'_i\alpha + (\rho\sqrt{2}\eta_i)}{\sqrt{1-\rho^2}}\right)$ . Once  $f(y_i | z_i, x_i, d_i, \sqrt{2}\sigma\eta_i)$  is specified, the log-likelihood function for a sample of size  $N$  can be established.

In the Poisson version of the model

$$f(y_i | z_i, x_i, d_i, \sqrt{2}\sigma\eta_i) = \frac{\{Exp\{x'_i\beta + \gamma d_i + \sqrt{2}\sigma\eta_i\}\}^{y_i} Exp\{-Exp\{x'_i\beta + \gamma d_i + \sqrt{2}\sigma\eta_i\}\}}{y_i!} \tag{4}$$

Notice that the mean and variance of the count variable are

$$\mu_i = E[y_i | d_i, z_i, x_i] = Exp\{x'_i\beta + \gamma d_i - 0.5\sigma\} \left( d_i \frac{\Phi(z'_i\alpha + \sigma\rho)}{\Phi(z'_i\alpha)} + (1-d_i) \frac{1-\Phi(z'_i\alpha + \sigma\rho)}{1-\Phi(z'_i\alpha)} \right) \tag{5}$$

and

$$Var(y_i|d_i, z_i, x_i) = \mu_i + k\mu_i^2 \tag{6}$$

where  $k = Exp(2\sigma^2) - Exp(\sigma^2)$ . Thus, the model exhibits overdispersion.

If it is assumed that there is no impact through the interaction effects with observed characteristics, and there are just effects due to being enrolled in the subsidized program and unobserved drivers because of self selection, the impact of the health care program can be measured by the following equation Maddala (1996).

$$\frac{E[y_i|z_i, x_i, d_i = 1]}{E[y_i|z_i, x_i, d_i = 0]} = Exp\{\gamma\} \left( \frac{\Phi(z'_i\alpha + \sigma\rho)}{\Phi(z'_i\alpha)} \right) \left( \frac{1 - \Phi(z'_i\alpha)}{1 - \Phi(z'_i\alpha + \sigma\rho)} \right) \tag{7}$$

where  $Exp\{\gamma\}$  measures the moral hazard effect and  $\left(\frac{\Phi(z'_i\alpha + \sigma\rho)}{\Phi(z'_i\alpha)}\right)\left(\frac{1 - \Phi(z'_i\alpha)}{1 - \Phi(z'_i\alpha + \sigma\rho)}\right)$  is the selection effect based on unobservable factors.

If  $\rho = 0$ ,  $e_i$  and  $v_i$  are independent, then  $d_i$  is an exogenous process, that is, there is no self selection under this formulation. Thus, the parameter estimates in an exogenous count model are unbiased and consistent. This model is identified through nonlinear functional form even if all the variables in the insurance equation are included in the utilization equation. However, we use exclusion restrictions to identify parameter estimates.

In the case of hospitalization, we follow the discussion of Amemiya (1978), Green (2003) and Miranda & Rabe-Hesketh (2006), where  $y_i = \{0, 1\}$  is a dichotomous variable which is equal to 1 if individual  $i$  was hospitalized in the last year. The model can be formulated as a system of equations for two latent variables ( $y_i^*$  and  $d_i^*$ ). The process for  $d_i$  is characterized by equation (1), where  $d_i^* = z'_i\alpha + v_i$  and  $y_i$  is generated by

$$y_i = \begin{cases} 1, & y_i^* = x'_i\beta + d_i\gamma + e_i > 0 \\ 0, & y_i^* = x'_i\beta + d_i\gamma + e_i \leq 0 \end{cases} \tag{8}$$

where  $v_i = \epsilon_i + u_i$  and  $e_i = \lambda\epsilon_i + \tau_i$ . Here,  $\epsilon_i$ ,  $u_i$  and  $\tau_i$  are distributed normal standard. Therefore,

$$\begin{bmatrix} e_i \\ v_i \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \lambda^2 + 1 & \lambda \\ \lambda & 2 \end{bmatrix} \right) \tag{9}$$

Thus,  $\rho = \frac{\lambda}{\sqrt{2(\lambda^2 + 1)}}$ .

If  $\lambda = 0$ , which implies  $\rho = 0$ , there is not self selection under this formulation, and the parameter estimates in a univariate Probit model are unbiased and consistent.

Again, the semi-structural model is identified, but we introduce exclusion restrictions to identify parameter estimates.

#### 4 Econometric Results

The subsidized health care program is targeted to poor households that accomplish some specific requirements. Specifically, they must be classified in

level one or two in the SISBEN,<sup>1</sup> and stratum one, two or three.<sup>2</sup> Additionally, all members of the household have to be unemployed.

We estimate our models using data from the Medellin Living Standards Survey in 2007. This is a cross-section representative survey of the non-institutionalized population in Medellin with excellent information on demographic and socioeconomic characteristics and health status. Out of a sample of 12,975 individuals that can freely choose to be in the subsidized program, 9,893 of them are in the program, that is 76.24% of the sample. As can be seen in Table 1, there is no statistical difference in the unconditional mean of the amount of preventive medical care utilization. The average number of visits is 2.45 and 2.49 for uncovered and covered individuals. On the other hand, the level of hospitalization of the former is 2%, while the latter is 3%. This implies that there is statistical difference in the level of hospitalization. In general, there are statistical differences between these groups. In particular, individuals involved in the program are older, there is a higher proportion of women, they have a lower level of education, their perception of health status is better in low levels and most of them are in stratum one and two.

Table 2 shows that 14.91% of the individuals did not visit a doctor for preventive purposes. Additionally, the cumulative relative frequency until four visits is 98.4%. The highest frequency is observed at four visits whose percentage is 36.31%. Probably, there is unconditional overdispersion in data because of the long right tail (the maximum number of preventive care visits is 64) and the unconditional variance is 3.63, which is greater than the mean that is 2.48.

We estimate count and binary models with endogenous switching where the former variable is the number of preventive medical care visits and the latter is a dummy indicator of hospitalization, both in the last year. The switching variable is to belong to the subsidized program. It must be taken into consideration that these semi-structural models are identified through non-linear functional forms, even if all the variables in the insurance equation are included in the utilization equations. However, given an initial specification with same controls in utilization and participation equations, specifically, we control by age, gender, level of education, own perceived health status and strata, we sequentially eliminate variables that were not statistically significant at 5% in each equation. Therefore, we use traditional exclusion restrictions to identify parameter estimates. Additionally, this procedure generates more efficient estimators.

As can be seen in Table 3, one of the most relevant characteristics in the Poisson count model is that there is a statistical significant correlation between unobserved factors that affect preventive health care utilization and subsidized program, as we can see that  $\rho$  is different from zero. Therefore, there is self selection, and individuals that hope to visit more frequently a doctor also show a higher probability of being in the program. Omitting this phenomenon implies biased and inconsistent parameter estimates. Additionally, we can observe from this table that there is evidence that overdispersion is present because of the fact that  $\sigma$  is statistically significant at 5%.

---

<sup>1</sup>SISBEN is a survey whose main objective is to identify and classify households that can obtain some governmental benefits.

<sup>2</sup>Households in Colombia are classified by strata. This is done in order to give subsidies associated with utility consumption to poor people who are classified in low strata. On the other hand, households that are classified in high strata must pay some contributions.

**Table 1:** Descriptive statistics

Variable	No Subsidized Regimen	Subsidized Regimen	Mean Hypothesis Test <sup>1</sup>
Number Preventive Visits	2.45 (2.17)	2.49 (1.82)	-0.99
Hospitalization	0.02 (0.16)	0.03 (0.18)	-2.42*
Age	25.96 (21.29)	32.37 (17.43)	-16.85*
Female	0.45 (0.50)	0.59 (0.49)	-13.56*
No Education	0.31 (0.46)	0.39 (0.49)	-7.94*
Primary	0.41 (0.49)	0.40 (0.49)	1.04
High School	0.25 (0.43)	0.20 (0.40)	5.82*
Vocational	0.01 (0.12)	0.01 (0.09)	2.49*
University	0.01 (0.06)	0.00 (0.11)	4.23*
Bad Health Status	0.01 (0.08)	0.02 (0.13)	-5.46*
Fair Health Status	0.07 (0.26)	0.13 (0.33)	-9.49*
Good Health Status	0.83 (0.38)	0.77 (0.42)	6.90*
Excellent Health Status	0.09 (0.29)	0.08 (0.28)	1.64
Stratum 1	0.24 (0.43)	0.26 (0.44)	-2.95*
Stratum 2	0.52 (0.50)	0.66 (0.47)	-13.80*
Stratum 3	0.24 (0.43)	0.08 (0.26)	20.41*

<sup>1</sup> Null hypothesis: means are equal. Critical value at 5% level of significance is 1.96.

\* Rejection of null hypothesis at 5% significance level.

Standard Deviation in parenthesis.

Source: Authors' estimations.

**Table 2:** Preventive Health Care Visits: Observed Frequencies

Medical appointments per year	Percent	Frequency
0	14.91	1935
1	21.87	2837
2	9.56	1241
3	15.75	2043
4	36.31	4711
More than 4	1.60	208

Source: Authors' calculations.



We can observe from Table 3 that older individuals visit more frequently a doctor but at a decreasing rate. Additionally, the number of visits are higher in people from stratum two compared with stratum one, but it is lower in stratum three. Individuals whose maximum level of education is primary do not have statistical difference compared with individuals without education, individuals whose education level is high school or university and that visit less frequently a doctor, but those whose maximum achievement education level is vocational do more preventive health care visits. Regarding self perceived health status, the base category is bad condition, and it can be seen from Table 3 that individuals whose health status is fair visit less frequently a health care center.

As we can see in tables 3 and 4, older women show a higher probability of being in the subsidized health care program, individuals in stratum two have a higher propensity compared with people in stratum one, on the other hand individuals in stratum three have a lower propensity. Regarding the level of education, the base category is no education, and we can observe that a higher education achievement implies a lower probability of being in the program. Finally, individuals whose self health perception is bad or fair show a higher probability of being enrolled.

It can be observed in Table 3 that moral hazard due to being in the subsidized program is  $-21\% = \text{Exp}\{-0.23\} - 1$ , and the mean self selection effect is 1.28 with a 95% confidence interval equal to (1.24, 1.37) (see equation 7). These values imply that individuals who are involved in the program visit a doctor 21% less frequently for preventive health care purposes than individuals who are not in this program. We apply the two-stage nonlinear method of moments developed by Terza (1998) as robustness check. We basically get the same outcomes with this methodology. However, it must be taken into consideration that this procedure is less efficient than Full Information Maximum Likelihood. See Annex, Table A.1.

In order to evaluate the effects of the subsidized health care program on hospitalization, we estimate an endogenous switching Probit model. As Table 4 illustrates, there is self selection due to  $\rho$  being statistically significant at 5%. Therefore, parameter estimators must take into consideration this fact to avoid bias and inconsistency problems. We can see in this table that individuals from stratum three and whose self health perception is bad have a higher rate of hospitalization.

Additionally, we can see in Table 4 that the subsidized health care program has a statistically significant effect on hospitalization. Specifically, there is a positive effect, which means that individuals enrolled in this program have a higher probability of hospitalization. For instance, a person enrolled in the program who is from stratum two and whose self perception of health status is good has a probability of 3.28%; whereas a person with these characteristics but who is not enrolled in the health subsidized program has a probability of 0.48%. We estimate an endogenous switching Logit model as robustness check, and we obtain similar results. See Annex, Table A.2.

## 5 Conclusions

Parameter estimates indicate that there is self selection, this fact implies that conventional techniques to evaluate public programs based on propensity

**Table 3:** Endogenous Poisson Model: Preventive Health Care

Variable	Coefficient	Standard Deviation	z Statistic
Number of visits			
Subsidized Regimen	-0.232	0.052	-4.480
Age	0.011	0.001	8.400
Squared Age	-0.0001	0.00001	-8.400
Stratum 2	0.044	0.015	2.920
Stratum 3	-0.167	0.027	-6.080
Primary	0.019	0.016	1.160
High School	-0.106	0.020	-5.290
Vocational	0.132	0.062	2.130
University	-0.197	0.087	-2.270
Fair Health Status	-0.113	0.057	-1.990
Good Health Status	-0.105	0.055	-1.920
Excellent Health Status	0.050	0.058	0.860
Constant	0.983	0.072	13.690
Subsidized Regimen			
Age*Female	0.014	0.001	20.610
Stratum 2	0.080	0.030	2.690
Stratum 3	-0.807	0.042	-19.310
Primary	-0.196	0.030	-6.650
High School	-0.267	0.034	-7.790
Vocational	-0.466	0.118	-3.970
University	-0.845	0.145	-5.850
Fair Health Status	-0.270	0.141	-1.910
Good Health Status	-0.453	0.136	-3.340
Excellent Health Status	-0.431	0.141	-3.050
Constant	1.143	0.137	8.320
$\sigma$	0.356	0.009	37.920
$\rho$	0.388	0.077	5.070

Source: Authors' estimations.

score matching obtain wrong outcomes, while endogenous models take into consideration this phenomenon, which implies unbiased and consistent estimates under the assumption of good statistical specification.

Regarding preventive health care utilization, it is found that people enrolled in this subsidized program visit less frequently a doctor compared with uncovered individuals. On the other hand, covered people have a higher probability of hospitalization than individuals who are not in this program. This result has bad consequences for financial statements of the subsidized health care program in Medellin due to hospitalization being more expensive than preventive services. Perhaps people involved in this program are less concerned about preventive services because they know that hospitalization services are not too expensive due to their situation. Therefore, our recommendation is to implement a strategy that encourages preventive health care utilization that mitigates possible negative effects on this program, and as a consequence to widen its coverage.

Regarding future research and limitations, we must take into consideration that although the endogenous Poisson model takes into account overdispersion, there is one limitation that has not been overtaken in this application, the excess of zeros. To overcome this characteristic it is necessary to implement a Zero Inflated Poisson model with endogenous switching. Although Greene (2008) develops a Zero Inflated Poisson model with selection, we did not find an econometric framework that takes into consideration endogenous

**Table 4:** Endogenous Switching Binomial Probit Model: Hospitalization

Variable	Coefficient	Standard Deviation	z Statistic
Hospitalization			
Subsidized Regimen	0.745	0.167	4.450
Stratum 2	-0.088	0.050	-1.750
Stratum 3	0.197	0.091	2.180
Fair Health Status	-0.470	0.112	-4.190
Good Health Status	-1.121	0.127	-8.830
Excellent Health Status	-0.993	0.144	-6.920
Constant	-1.376	0.181	-7.610
Subsidized Regimen			
Age*Female	0.014	0.001	20.350
Stratum 2	0.080	0.030	2.710
Stratum 3	-0.809	0.042	-19.490
Primary	-0.192	0.029	-6.550
High School	-0.263	0.034	-7.680
Vocational	-0.438	0.122	-3.590
University	-0.857	0.144	-5.960
Fair Health Status	-0.237	0.138	-1.720
Good Health Status	-0.421	0.132	-3.200
Excellent Health Status	-0.400	0.138	-2.910
Constant	1.110	0.133	8.340
$\rho$	-0.450	0.103	-4.360

Source: Authors' estimations.

switching.

### Acknowledgements

We thank participants of Banco de la República seminar for fruitful comments and discussions. However, all remaining errors are our own.

### Bibliography

Amemiya, T. (1978), 'The estimation of a simultaneous equation generalized probit model', *Econometrica* **46**(5), 1193–1205.

Cameron, A. C. & Trivedi, P. (2005), *Microeconometrics: Methods and Applications*, first edn, Cambridge University Press.

Cameron, A., Trivedi, P., Milne, F. & Piggot, J. (1988), 'A microeconomic model of the demand for health care and health insurance in australia', *Review of Economic Studies* **55**, 85–106.

Congress, C. (1993), Law 100 of 1993.

**URL:** <<http://www.caprecom.gov.co/sitio/filesnormatividad/Ley-100-de-1993.pdf>> Accessed May-2010

Gaviria, A., Medina, C. & Mejia, C. (2006), Evaluating the impact of health care reform in colombia: From theory to practice, Working Paper 1657-7191, Universidad de Los Andes.

Green, W. (2003), *Econometric Analysis*, fifth edn, Prentice-Hall.

Greene, W. (2008), 'Models for count data with endogenous participation', *Empirical Economics* **36**, 133–173.

Maddala, G. S. (1996), *Limited-Dependent and qualitative variables in econometrics*, Cambridge University Press.

Miranda, A. (2004), 'Fiml estimation of an endogenous switching model for count data', *The Satat Journal* **4**(1), 40–49.

Miranda, A. & Rabe-Hesketch, S. (2006), 'Maximum likelihood estimation of endogenous switching and sample selection models for binary, ordinal and countvariables', *The stata journal* **6**(3), 285–308.

Pita, P., Machado, M. & Sanz, A. (2008), 'Moral hazard and the demand for health services: A matching estimator approach', *Journal of health economics* **27**(4), 1006–1025.

Saadah, F., Pradhan, M. & Sparrow, R. (2001), The effectiveness of the health card as an instrument to ensure access to medical care for the poor during the crisis, Working paper, World Bank.

Terza, J. (1998), 'Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects', *Journal of Econometrics* **84**, 129–154.

Trujillo, A., Portillo, J. & Vernon, J. (2005), 'The impact of subsidized health insurance for the poor: Evaluating the colombian experience using propensity score matching', *International Journal of Health Care Finance and Economics* **5**, 211–239.

Urdinola, D. & Jain, S. (2006), Don subsidized health programs in armenia increase utilization among the poor?, Working Paper 4017, World Bank.

Wagstaff, A. (2007), Health insurance for the poor: Initial impacts of vietnam's health care fund for the poor, Working paper 4134, World Bank.

Wagstaff, A., Lindelow, M., Jun, G., Ling, X. & Juncheng, Q. (2007), Extending health insurance to the rural population: An impact evaluation of china's new cooperative medical scheme, Working paper 4150, World Bank.

Appendix A

**Table A.1:** Two Stage Method of Moments: Preventive Health Care

Variable	Coefficient	Standard Deviation	z Statistic
Number of visits			
Subsidized Regimen	-0.168	0.141	-1.190
Age	0.010	0.001	7.390
Squared Age	-0.0001	0.0000	-7.260
Stratum 2	0.042	0.016	2.690
Stratum 3	-0.136	0.047	-2.890
Primary	0.023	0.018	1.290
High School	-0.097	0.024	-4.080
Vocational	0.135	0.069	1.950
University	-0.177	0.080	-2.210
Fair Health Status	-0.150	0.100	-1.500
Good Health Status	-0.146	0.099	-1.470
Excellent Health Status	0.002	0.100	0.020
Constant	1.038	0.163	6.360
Subsidized Regimen			
Age*Female	0.014	0.001	20.460
Stratum 2	0.081	0.030	2.730
Stratum 3	-0.809	0.042	-19.340
Primary	-0.193	0.030	-6.530
High School	-0.270	0.034	-7.870
Vocational	-0.450	0.119	-3.790
University	-0.846	0.145	-5.830
Fair Health Status	-0.269	0.142	-1.890
Good Health Status	-0.445	0.137	-3.250
Excellent Health Status	-0.424	0.142	-2.980
Constant	1.136	0.138	8.210
$\theta$	0.099	0.086	1.150

Source: Authors' estimations.

**Table A.2:** Endogenous Switching Binomial Logit Model: Hospitalization

Variable	Coefficient	Standard Deviation	z Statistic
Hospitalization			
Subsidized Regimen	0.860	0.066	12.940
Stratum 2	-0.099	0.059	-1.670
Stratum 3	0.220	0.092	2.400
Fair Health Status	-0.479	0.155	-3.090
Good Health Status	-1.191	0.337	-3.530
Excellent Health Status	-1.040	-0.310	-3.360
Constant	-1.483	0.152	-9.770
Subsidized Regimen			
Age*Female	0.014	0.001	20.250
Stratum 2	0.080	0.030	2.710
Stratum 3	-0.809	0.042	-19.480
Primary	-0.192	0.029	-6.540
High School	-0.263	0.034	-7.630
Vocational	-0.437	0.122	-3.570
University	-0.857	0.144	-5.970
Fair Health Status	-0.233	0.137	-1.700
Good Health Status	-0.418	0.131	-3.200
Excellent Health Status	-0.398	0.137	-2.900
Constant	1.108	0.133	8.360
$\rho$	-0.498	0.103	-4.850

Source: Authors' estimations.