International Journal of Health Care Finance and Economics, 5, 211–239, 2005

The Impact of Subsidized Health Insurance for the Poor: Evaluating the Colombian Experience Using Propensity Score Matching

ANTONIO J. TRUJILLO* atrujill@mail.ucf.edu

College of Health and Public Affairs, University of Central Florida, Orlando

JORGE E. PORTILLO

Economists Incorporated, Washington DC

JOHN A. VERNON

School of Business, University of Connecticut

This paper evaluates the impact of Colombia's subsidized health insurance program (SUBS) on medical care utilization. Colombia's SUBS program is a demand-side subsidy intended for low-income families, where the screening of beneficiaries takes place in decentralized locations across the country. Due to the self-selection problems associated with non-experimental data, we implement Propensity Score Matching (PSM) methods to measure the impact of this subsidy on medical care utilization. By combining unique household survey data with community and regional data, we are able to compute propensity scores in a way that is consistent with both the local government's decision to offer the subsidy, and with the individual's decision to accept the subsidy. Although the application of PSM using these rich datasets helps to achieve a balance between the treatment and control groups along observable dimensions, we also present instrumental variable estimates to control for the potential endogeneity of program participation. Using both methods, we find that Colombia's subsidized insurance program greatly increased medical care utilization among the country's poor and uninsured. This evidence supports the case for other Latin American countries implementing similar subsidy programs for health insurance for the poor.

Keywords: medical care use, health insurance program evaluation, propensity score, colombia

JEL classification:

1. Introduction

Subsidized health insurance programs are often used by developing countries to provide basic health care to their poor and uninsured citizens. These programs have been met with

*Address correspondence to: 14140 Cheval Vineyard Way, Apt. 105, Orlando, FL 32828, USA. Tel.: (407) 823-4238; Fax: (407) 823-6138.

varying degrees of success. The main justification for this type of transfer is the belief that adequate medical care is a fundamental human right. The potential distributional gains from such subsidies are, however, usually accompanied by efficiency losses, which occur because of distortions in both the participants' behavior and in the allocation of resources to the health care sector as a whole. As a result, it is imperative for policy makers to carefully evaluate the effectiveness of these programs, both in terms of their allocative inefficiencies and in terms of the extent to which they achieve their desired distributional objectives. Unfortunately, comprehensive evaluations of subsidized health insurance programs in developing countries are rarely undertaken.

This paper evaluates the impact of Colombia's subsidized health insurance program, henceforth SUBS, on the level of medical care utilization. In Latin America, medical care is typically provided by public institutions and paid for by governmental single payer institutions. Colombia is an exception: faced with efficiency, equity, and quality problems in the delivery of medical services, in 1993 its government decided to profoundly reform the state-dominated health care sector. It chose a regulated competition approach, increasing the participation of private health insurers in a way that is, to date, unique in the region. Competition, reformers hoped, would lower health care costs and increase quality, while state involvement ensured equity. If these reforms succeed at introducing market rationality and preserve the solidarity that characterized the old system, they will likely be imitated. Specifically, the SUBS program is a demand-side government subsidy intended for low-income families to buy health insurance. The screening of beneficiaries is decentralized and conducted by local government officials.

In addition to addressing and evaluating this policy initiative, our paper also focuses on the broader issue of applying non-experimental methods to evaluate social programs in the health care sector. Due to the lack of randomized study data in developing countries on the impact of health insurance on medical care consumption, the method proposed in this paper is Propensity Score Matching (PSM), a method recommended by the World Bank for the evaluation of social programs such as Colombia's SUBS program (Baker, 2000). However, as a robustness check, and to control for potential endogeneity, we also employ instrumental variable estimation techniques, and compare these findings with our PSM results.

It is well known that private health insurance coverage significantly increases medical care utilization because of moral hazard (Pauly, 1968). However, in the case of public health insurance programs, especially in developing countries, less is known about the magnitude of the real effect of insurance coverage on medical care use. For instance, individuals outside the insurance program could be in worse health than individuals in the program. Sickness funds could have an incentive to "cherry picking" or "skimming the cream" of subscriber pool in order to maximize profit. This behavior may obscure the true effect of subsidizing health insurance programs for the poor.

Due to the self-selection problems associated with non-experimental data, comparing medical care use between program participants (the treatment group) and non-participants (the control group) will result in a biased estimate of the program's effect. This is because both groups are different in terms of their individual characteristics–characteristics that may influence *both* their level of medical care utilization and their decision to participate in the program. With experimental data, this bias disappears because the random assignment of individuals to each group balances the observable and unobservable individual characteristics affecting medical care use (Joffe and Rosenbaum, 1999). Although experimental data reduce bias due to self-selection problems, it does not do so without some costs. For example, it may be either prohibitively expensive to conduct a randomized study, or it may be unethical. Heckman and Smith (1995) provide an excellent review of the literature on the limitations of experimental design in social science research. For this reason, considerable effort has been devoted towards improving the efficiency of non-experimental methodologies in estimating the average impact of new social programs.

In the absence of experimental data, researchers may try to replicate the features of random allocation by using observational data and standard multiple regression analyses. However, this requires strong assumptions about the distribution of the unobservables, and a priori information on the functional form and relationship between, for example, medical care consumption and its principal determinants. PSM is an alternative approach to assessing average program effects, and it has gained much popularity in recent years (Rosenbaum and Rubin, 1983, 1984, 1985; Dehejia and Wahba, 1999; Smith and Todd, 2001). Specifically, PSM balances the observable variables in a cross-sectional database between the treatment and control groups. Assuming that unobservable pre-treatment differences between the two groups do not exist, any difference in the outcome of interest will be due exclusively to the program effect. Unfortunately, PSM does not balance the *unobservable* variables between groups (Joffe and Rosenbaum, 1999). Despite this distributional assumption, however, PSM offers two practical advantages to the program evaluator. First, it provides a measure of the extent to which common characteristics in the treatment and control groups differ (Dehejia and Wahba, 1999). Second, as previously mentioned, PSM allows the program evaluator to estimate the effect of the program without explicitly modeling the relationship between the individual characteristics and the outcome of interest. This may be particularly advantageous when evaluating the Colombia SUBS program because modeling the consumption of medical care use for Colombia's poor and uninsured would be an extremely complex task. Moreover, for the evaluator of the program, it is only the program's effect on medical care utilization that is of primary interest. It is important to note, however, that other methodologies, such as instrumental variables estimation (IVs), are not always possible, and like PSM they also impose restrictive assumptions about the distribution of the unobservable covariates. While the literature is rich with debate on the merits of the various non-experimental methodologies, for social program evaluations such as Colombia's SUBS, we believe we are on firm ground in utilizing a PSM methodology: a considerable precedent has been established in the labor economics and statistics literatures (Heckman and Hotz, 1989; Dehejia and Wahba, 1999; Sianesi, 2001).

This being said, however, even under the PSM distributional assumption, two factors have been suggested in the literature (see Smith and Todd, 2001 for a recent development) that may compromise the performance of PSM estimates: the quality of the data and the ability to impose a common support condition. In our case, combining household, community, and regional data allow us to overcome these limitations. In fact, we computed the propensity score controlling for an atypically rich set of relevant variables within homogenous age groups, and were able to restrict the matching process to an area of common support. We then estimated the impact of Colombia's SUBS by comparing those in the subsidized regime (treatment group) with those without any kind of health insurance (control group). We used three separate outcome indicators: preventive care, outpatient care, and hospitalization. Separate estimations were undertaken for various age groups. Finally, in order to ensure the robustness of our results, we apply four different criteria to determine the optimal match: stratification, nearest neighbor, radius matching, and the kernel method.

Despite the fact that using PSM with these rich datasets helps us to achieve a good balance between treatment and control groups along observable dimensions, one could argue that relevant unobservable confounding differences may still exist between groups. If this is the case, PSM estimates of program participation would be biased. We therefore complement our analysis by presenting instrumental variable estimates to control for these potential unobserved differences. Although the two methods are not directly comparable, we believe that presenting IV estimates along with our PSM results will generate a more robust finding, and will serve as a robustness check with respect to the effect of subsidized health insurance program on medical care use. It is important to keep in mind, however, that in PSM one is matching observed characteristics, whereas in IV estimation theoretically justified exclusion restrictions are used. In this latter method, variability in the exclusion restriction is used to estimate the effect, rather than similarity in the "exclusion restriction" to make observations close.

Our paper will proceed as follows. Section 2 provides a detailed overview of the institutional characteristics and background of Colombia's SUBS program. Section 3 then presents our evaluation framework and the PSM methodology. In Section 4 we outline and discuss our overall empirical strategy, including our IV estimation strategy. Section 5 follows and presents and discusses our results. Section 6 concludes the paper.

2. Institutional Background

2.1. The Health Insurance Subsidy Program (SUBS)

In this section, we will focus on two important aspects of the SUBS program: its general design characteristics, and its eligibility criteria. In response to efficiency and equity concerns regarding the delivery of medical services, the Colombian government initiated (in the early 1990s) a profound reform of the country's health care sector. The reform separates the provision of health insurance from the delivery of medical care, minimizes the government's role in the industry, decentralizes assistance programs to local governments, and introduces a social insurance system based on regulated competition among private firms.

By the end of 1995, most of the legislation had been implemented, resulting in a system with three tiers: (i) a contributory social insurance regime financed by mandatory payroll taxes; (ii) a subsidized regime that targets low-income and disadvantaged groups and that, for the most part, is financed with general taxes; and (iii) a publicly-financed safety net that provides basic medical services for the uninsured. In addition, there is a private, premium-based health insurance market for those who want and can afford the extra coverage. For a detailed description of the Colombian health care sector before and after the health care reform, see Bloom, Mina and Olivo (1998), Velez and Foster (2000), and Peñaloza and Londoño (2002).

The subsidized regime is financed by general taxes plus a constant fraction of the payments made to the contributory regime, including the special insurance schemes. This funding is transferred to local governments, who allocate it, within certain boundaries, between supply and demand subsidies. On the supply side, a percentage of the transfers earmarked for health should go to subsidize the operations of autonomous, non-for-profit medical care providers known as ESS (*Empresas Solidarias de Salud*). On the demand side, local elected officials, usually the mayor, are responsible for identifying eligible individuals and enrolling them into special non-for-profit insurance companies known as ARS (*Administradoras del Régimen Subsidiado*). Insurers contract medical services with providers and compete on their ability to offer the government-defined health package at the lowest costs. In order to control an excessive use of medical care services, patients enrolled in SUBS pay a coinsurance rate that vary between 5% and 30% according to the individual's income.

Overall, the Colombian health care reform has achieved considerable success in increasing insurance coverage, particularly among the lower income groups. Thus, the proportion of the population with some form of health insurance doubled during the 1993–1997 period and, as reported by Londoño, Jaramillo and Uribe (2001), the largest proportional gains were registered in the lower income quintiles.[1] Although much of these distributive gains respond to the progressive implementation of the subsidized regime, an important proportion of low-income individuals do not receive the subsidy. Even more troublesome, is the fact that a non-trivial proportion of individuals at the highest levels of income do receive the subsidy, which bring us to the issue of targeting.

2.2. Targeting of the Subsidized Regime

The SUBS program targets its intended population through a system that combines individual means-testing with elements of both categorical targeting (tagging) and self-targeting. Specifically, the main tool for determining eligibility into the regime is a proxy means-test known as SISBEN (*Sistema de Selectión de Beneficiarios para Programas Sociales*). This is a crude welfare index that ranks families according to a set of household characteristics, human capital endowment, and reported income level. Municipal governments are responsible for conducting a survey among the poorest neighborhoods in the county, using a short, standardized questionnaire.[2] All the surveyed families are then ranked into six levels and, in principle, only families in levels 1 and 2 (the poorest) are eligible for subsidized health insurance. Those in level 3 may receive the subsidy only if funding is available and those in the lowest two levels have been taken care of.

Because there is no guarantee that local governments will receive enough funding to grant the subsidy to all eligible individuals, the subsidized regime gives priority to certain groups defined by easily identifiable characteristics or tags. Specifically, within the pool of individuals at low SISBEN levels, municipal authorities should give priority to rural and indigenous populations, pregnant women and children under five, handicapped individuals, senior citizens, and female-headed households.

Although participation in the contributory regime is mandatory for all workers earning more than twice the minimum wage, it is not uncommon to find individuals who get around such regulation, particularly those who are self-employed. As a disincentive to this kind of "affiliation swap", the SUBS imposes significant non-pecuniary costs (waiting, red tape) on its affiliates and offers a health care package that is less comprehensive than the one available to those in the contributory regime.

As indicated in the previous section, a significant number of low-income families are being excluded from the subsidized regime while others at the highest levels of income do receive the subsidy. These errors of exclusion and inclusion in SUBS are the result of three major sources of distortion.[3]

First, due in part to imperfect monitoring, some local officials misdirect or misuse program resources (at least, by the center's criteria). In the case of SUBS, the central government delegates responsibility for candidate selection to municipal governments. In doing so, the center defines the selection criteria, provides technical support, and, at least in theory, guarantees a capitated payment for every eligible individual identified. In exchange, municipal officials are expected to use their superior information about local needs to optimize the selection process. In practice, the central government may only guarantee funding for a fraction of those eligible, in which case municipal officials are in charge of rationing the subsidy. There is anecdotal evidence that some officials take advantage of their discretionary power and bend the selection process for political gain (Jaramillo, 2001), but it is also true that others are simply responding to welfare criteria that, although different from the center's, are not devoid of merit. For example, local officials may prefer to enroll a household facing a grave illness, even if they are not poor enough to qualify, rather than giving the subsidy to a poor but relatively healthy household (McGee, 1999).

Second, some municipal governments face very limited administrative capabilities to identify and enroll eligible individuals. For example, Bloom, Mina and Olivo (1998) found that, after controlling for the incidence of poverty and the urban-rural make up of the population, larger municipalities show a larger SUBS affiliation rate, in terms of both population and total fiscal transfers from the central government. Basically, smaller municipalities cannot fully exploit the scale economies inherent to the survey implementation and related fieldwork. These municipalities may find it attractive to devote a relatively larger share of their health budget to supply subsidies, which, although less focalized, are easier to manage.

Finally, even if local governments had adequate administrative capabilities and would not deviate from the center's welfare criteria, deficiencies inherent to the proxy means-test used result in inclusion and exclusion errors. Most notably, the items in the SISBEN questionnaire regarding employment status and reported income yield a poor measure of household income in general and do not reflect rural working conditions in particular (McGee, 1999). This and other shortcomings have been well-documented (Velez, Castano and Deutsch, 1999) and, according to Colombia's Ministry of Planning (DNP).[4]

2.3. Previous Studies About the Subsidized Regime

There are several empirical studies that evaluate the Colombian health sector after the reform, including the performance of the subsidized regime. Nevertheless, most of these studies are descriptive in their nature and rely mostly on aggregated data. See for example: Bloom, Mina and Olivo (1998), Londoño, Jaramillo and Uribe (2001), Peñaloza and Londoño (2002), and Bitran et al. (2003).

A notable exception is a study by Panopoulus and Velez (2002) that uses household level data to explore the determinants of enrollment in the SUBS program and its impact on medical care use. Panopoulus and Velez used the 1997 Colombia Living Standards Survey (ECV/97),[5] to estimate their own approximation to the SISBEN index, and then focus their attention on the poorest households (levels 1 and 2). For this sub-sample, they estimate an equation of medical care use and find that, after controlling for various supply and demand factors, having a SISBEN classification document has little effect on medical care use.[6] This unexpected result should be taken with caution since this study has several limitations that make its finding difficult to extrapolate. The main limitations include but not limited to: (i) it limits the analysis to the head of the household; (ii) it wrongly assumes that the SISBEN index is the only criterion used to select beneficiaries of SUBS; (iii) more importantly, the study does not evaluate whether or not medical care use has been affected by participation in the subsidized regime per se. In this research, we overcome most of the limitations of previous empirical attempts to evaluate the impact of subsidized health insurance for the poor in Colombia. Faced with efficiency, equity, and quality problems in the delivery of medical services, the Colombian government decided to reform the state-dominated health care sector in 1993. It chose a regulated competition approach, increasing the participation of private health insurers in a way that is, to date, unique in the region. If these reforms succeed at introducing market rationality in the subsidized system and preserve the solidarity that characterized the old system, they will likely be imitated. Results from this analysis would provide insight for improving implementation of similar health programs in the region.

3. Program Evaluation Framework

Following a common notation in the program evaluation literature (e.g., Heckman, Ichimura and Todd, 1997; Dehejia and Wahba, 1999; Smith and Todd, 2001; Minkin, 2002), let Y_{i1} indicate the value of the outcome variable for a participant (treatment group), while Y_{i0} denotes the value for the non-participant (control group). The expected treatment effect (ET) is

$$\Delta = E(Y_{i1}) - E(Y_{i0}). \tag{1}$$

However, due to self-selection problems in non-experimental data, the comparison of participants and non-participants in (1) would result in biased estimates of the program effect. Thus, the relevant question for a policy analyst is what would have happened with the participants if they had not participated in the program, or what is called the *expected treatment effect for the treated* (ETT). Defining $W = 1$ for participants and $W = 0$ for non-participants, ETT is

$$\Delta|_{W=1} = E(Y_{i1}|Wi = 1) - E(Y_{i0}|Wi = 1). \tag{2}$$

In experimental data, the bias disappears because random allocations balance the observable and non-observable individual characteristics among participants and non-participants, and equation (1) gives an unbiased ETT. Formally, random allocation implies that Y_{i0},

$Y_{i1} \perp W_i$, and hence $E(Y_{i0}|Wi = 1) = E(Y_{i0}|Wi = 0)$. In this case,

$$\Delta|_{W=1} = E(Y_{i1}|Wi = 1) - E(Y_{i0}|Wi = 1) = E(Y_{i1}|Wi = 1) - E(Y_{i0}|Wi = 0) \quad (3)$$

Note that in non-experimental data this last equation does not hold and researchers do not have the data to compute $E(Y_{i0}|Wi = 1)$. Different methodologies and assumptions need to be made to solve this missing data problem. One of these approaches is cross sectional propensity score matching (PSM).[7] This approach is of particular relevance when there is no baseline for comparison and only non-experimental data is available.

The propensity score is the conditional probability of participating in the program given certain observable variables (Rosenbaum and Rubin, 1983). Given a vector of observable variables X, the propensity score is $z(X) = \text{prob}(W = 1|X) = E(W|X)$. A propensity score has three important properties: (a) it balances observable variables; (b) if adjusting for observable variables, X removes differences between the control and treatment groups and it is appropriate to adjust for their propensity score; (c) using an estimated propensity score to remove bias is better than using X since the estimated propensity score removes non-systematic differences in X.[8]

Formally, properties (a) and (b) imply

$$Y_{i1}, Y_{i0} \perp Wi|(X_i). \Rightarrow Y_{i1}, Y_{i0} \perp Wi \,|\, z(X_i) \quad \forall i. \quad (4)$$

Applying (4) to equation (2), we obtain that ETT equals:

$$\begin{aligned} \Delta \,|\, _{W=1} &= E(Y_{i1} \,|\, z(Xi), Wi = 1) - E(Y_{i0} \,|\, z(Xi), Wi = 1) \\ &= E(Y_{i1}) \,|\, z(Xi), Wi = 1) - E(Y_{i0} \,|\, z(Xi), Wi = 0). \end{aligned} \quad (5)$$

Using propensity score $E(Y_{i0}|z(Xi), Wi = 1)$ can be approximated by $E(Y_{\text{i}0}|z(Xi), Wi = 0)$. Thus, assuming that unobservable pre-treatment differences among participants and non-participant do not exist, the PSM based on a full set of observable variables X provides an unbiased estimated of ETT. Note that ETT could be estimated without modeling the relationship between X and the outcome of interest. Also, notice that we are interested in the propensity scores; therefore, we do not need to recover the unbiased structural parameters of the determinants of individual's participation.

4. Empirical Strategy

4.1. Specification of the Propensity Score Function

To simplify matters, we propose a conceptual framework where the SUBS status of an individual depends on two sequential decisions: the local government's decision to offer the benefit and the individual's decision to take it or actively look for it. On the supply side, although the targeting system SISBEN was intended to work almost as a mathematical formula, local officials clearly have some leeway during its implementation. Thus, the government's decision to offer the benefits g(·) can be viewed as a function of its administrative capabilities, the political preferences of the local authorities, the extend of their grassroots support, and the degree of public scrutiny.

On the demand side, the individual's decision to accept or seek the subsidy t(·) is a function of the individual net benefits associated with the program. Specifically, the benefits of enrolling in the program would depend on the health status of the beneficiary, her preferences toward medical care, her attitudes toward financial risk, and the level of access to medical care (Remler and Glied, 2003). The potential benefits have to be weighted against the financial and non-financial costs associated with enrolling into the program, including the cost of searching for information. Moreover, some analysts have suggested that receiving certain government benefits may create inconvenience for the beneficiary due to cultural factors, or the stigma of participating in government assistance programs (see for example Moffitt, 1983; Storer and Van Audenrode, 1995).

Since we are not interested in recovering structural parameters of the participation equation, the specification of $z(X)$ can be treated as a reduced form model that contains the arguments in $g(\cdot)$ and $t(\cdot)$.

4.2. Estimation of the Propensity Score

Initially, we estimate the propensity score $z(X)$ using a logit model that collapses the arguments of both the $g(\cdot)$ and the $t(\cdot)$ functions. Nevertheless, an incorrect specification of the propensity score may lead to biased estimates of ETT because the distribution of the pre-determined observable variables would not be balanced between the treatment and the control group (i.e., balancing property). To overcome this problem, we use an algorithm that tests a necessary condition for the balancing property[9] (See Becker and Ichino, 2002). The rich data sets used in this analysis allow us to calibrate the $z(X)$ until the balancing property function was achieved. Clearly, we could not test the *ignorable treatment assignment property*. However, in practice a correctly specified propensity score function provides crucial information to detect how similar are the controls and treated, in terms of mean values of the observable covariates.

4.3. Matching Estimators

From the previous step, we obtain blocks with different numbers of treatment and control units, making sure that within blocks, the average propensity score between treated and controls are equal, and that the average for each explanatory variable are also equal. Note that, since $z(X)$ is a continuous variable, there are not two observations with exactly the same estimated score. Now, the relevant question becomes how to decide which observations are close matches. In order to ensure the robustness of our results, we apply four different criteria to determine the optimal match: the stratification method, the nearest neighbor, the radius matching, and the kernel method. Becker and Ichino (2002) provide details regarding the algorithm used to compute each matching estimator, the pros and cons of each method, and the different numerical conditions imposed in the estimations. In general, the quality of the estimates can be improved if each matching estimator is calculated where the average propensity score overlap for the treated and controls, i.e., the common support condition.

4.4. Instrumental Variable Estimates

One may argue that PSM does not completely surmount the endogeneity of SUBS participation. Basically, after balancing the treatment and the control groups along observable dimensions, it is possible that unobserved confounding variables still remain between the two groups. If this is the case, the assumption of ignorable treatment does not hold, and PSM will result in biased estimates of the program effect. Although we believe that the richness of our data sets make the assumptions behind PSM plausible, we decided to complement our analysis of the SUBS effect on medical care use by implementing IV's estimates.

We use two-stage least squared regression (2SLS) models with exclusion restrictions that help us explain variation in program participation that is independent of an individual's decision to use medical care. When one uses a single equation of medical care utilization, failure to control for individual health status or individual preference toward medical care biases the relevant parameter estimates of the demand for medical care. Alternatively, in this analysis we estimate a first equation of participation as a function of individual characteristics, household characteristics, community characteristics and the instrumental variables (i.e., exclusion restrictions). The estimated individuals' participation is then used in a second equation to estimate the effect of the subsidized program on medical care use after controlling for a vector of individual, household, and regional and community characteristics. If the IV's are correct, the 2SLS method will generate unbiased estimates of the program effect on medical care consumption.

We construct different instrumental variables using aggregate community information. In general, these IV's should have three desirable characteristics: they should explain a high variation of the individual's decision to participate in the program; they should not be correlated with the individual's decision to use medical care; and finally, they should be uncorrelated with each other and with the rest of the control variables used in the participation and medical care equations. Among the variables that we considered as IV's are: health center in the community (Red de Solidaridad Social), the existence of government-sponsored grassroots organizations in the community (*Red de Solaridad Social*), Municipal Living Standard Index for 1993, voter turn out in 1994 municipal elections, and percentage of state population affiliated to the subsidized regimen. From a conceptual point of view, all these variables may explain the individual's decision to participate in the program and theoretically should be unrelated to the individual's decision to use medical care. Furthermore, given that the level of aggregation is at the community level, one may expect that these variables are uncorrelated with the rest of the control variables used in the analysis.

We run the 2SLS regression analysis for the three measures of medical care consumption available in the data set: preventive care, outpatient care and hospital use. For the purpose of contrasting results, we estimate the model for the same age groups as used in the PSM analysis. Finally, the validity of the IV estimation was evaluated using the standard statistical test developed by Davidson and MacKinnon (1993) where the predicted value of individual's participation is added as a right-hand-side variable in the estimation of the medical care equation. One should notice that this augmented estimation of the medical care equation would include the original insurance information with the estimated information about individual's participation. Since there is only one endogeneous variable (i.e., SUBS participation); this test is equivalent to a simple t-test for the predicted value of individual's participation in the medical care equation.

4.5. Data Description

A potential downside of the PSM method is that it requires a significant amount of data to balance observable between the treated and controls. In fact, with small data sets, researchers may face the dilemma of matching observations with close propensity score but with important differences in individual characteristics, or having only few units to match within blocks. In our case, we address this concern by using the 1997 Colombia Living Standards Survey (ECV/97). This is a multitopic, nationally representative survey covering 29 of the 33 states in the country. Originally designed for 10,000 households, the final tally was 9,121 households, with a total of 38,518 individuals. This household level information is complemented by a community questionnaire for the corresponding 621 urban communities and 397 rural communities. By using such a rich dataset, we are able to fully specify the participation equation, meet the balancing property, and impose the common support condition.

Unfortunately, the ECV/97's household questionnaire was drafted in a way that, for some observations, makes it difficult for us to clearly establish affiliation to the subsidized regime, an issue that Panopoulus and Velez also has to grapple with. Although they acknowledge that possession of an SCD is not equivalent to affiliation into SUBS, they opt to use the former as a perfect identifier of the treatment and control groups. We take a more conservative approach and define as affiliated only those individuals that: (i) are in possession of a SCD; (ii) are enrolled in any insurance company (ARS, ESS, EPS, or *Caja*) that is part the social health insurance system;[10] (iii) do not make any monthly insurance payments and no contribution is made by their employer or a relative; and (iv) do not have any private or complementary health insurance. Individuals that meet all these conditions constitute our treatment group. By contrast, our control group is formed by those individuals without an SCD and without any form of health insurance.[11] As a result, we end with a sample of 22,291 individuals, 5,559 in the treatment and 16,732 in the control group.

We make separate estimations for three dichotomous variables of medical care use: (i) whether or not the individual has used **preventive** care in the 12 months prior to the interview, (ii) whether or not she has been **hospitalized** in the 12 months prior to the interview, and (iii) whether or not she has used **outpatient** care in the 30 days prior to the interview. Table 1 reports the unconditional means of medical care use between participants and non-participants. For the most part, participants of all ages have higher medical care use than non-participants, with the exception of preventive care access for individuals 60–98 years. However, one should keep in mind that these differences do not represent unbiased estimates of the ETT since participants may differ from non-participants in relevant personal characteristics.

Table 1. Groups differences in unconditional means of medical care use estimated expected treatment effect (ET) ($N = 22{,}291$).

Age/Group	Preventive	Outpatient	Hospital	Sample size
Age 0–4				
Treatment	0.374	0.197	0.071	640
Control	0.351	0.203	0.049	2,064
Difference	0.023	−0.006	0.022**	
Age 5–15				
Treatment	0.303	0.132	0.025	1,701
Control	0.291	0.113	0.025	4,302
Difference	0.012	0.020**	0.000	
Males age 16–59				
Treatment	0.254	0.158	0.054	1,356
Control	0.272	0.121	0.043	4,913
Difference	−0.018	0.037***	0.011*	
Females age 16–59				
Treatment	0.340	0.283	0.112	1,505
Control	0.367	0.230	0.089	4,666
Difference	−0.027*	0.053***	0.023**	
Age 60–98				
Treatment	0.286	0.281	0.098	357
Control	0.321	0.247	0.072	787
Difference	−0.035	0.035	0.026*	
Rural				
Treatment	0.274	0.167	0.059	3,437
Control	0.271	0.136	0.046	8,453
Difference	0.003	0.031***	0.013***	
Urban				
Treatment	0.363	0.241	0.076	2,122
Control	0.361	0.195	0.059	8,279
Difference	0.002	0.046***	0.017***	

***Significant at $p < 0.01$, **Significant at $p < 0.05$, *Significant at $p < 0.10$

The independent variables that we use to explain the probability of affiliation into SUBS can be grouped in three major vectors: individual, household, and local characteristics. Since our objective is not to estimate a structural model of affiliation, these three vectors contain variables that affect both the individual's decision to participate in SUBS and the government's decision to offer the program.

The first vector of variables includes individual's health related variables (age, gender, self-reported health status, and existence of a chronic condition) and individual's socio-economic characteristics (employment status, sector of employment, education level, marital status). The health related variables will capture the individual preferences toward medical care use and her expected benefits from participating in the program. For instance, individuals in poor health conditions may assign a higher value to the program than individuals in good health conditions. The socio-demographic characteristics will include elements that affect both the costs and benefits of participating in the program. For instance, marital status and family size are considered elements that affect family availability of resources to produce health as well as preferences toward the program. While education level may affect the costs to enroll in the program since individuals with higher education will face lower costs of searching for information related to the program. Alternatively, one could argue that higher level of education increase preference toward medical care in order to maintain the individual's stock of health.

For those under 16 years of age, we hypothesize that the probability of participating in the program is associated with how the head of the household perceives the costs and benefits of enrollment. Therefore, for these groups, individual characteristics like education and employment status refers to the head of the household and not to the minor.

A second vector of covariates corresponds to household characteristics (such as household size, number of rooms, percentage of household members with a job, and housing characteristics) that try to capture the individual's eligibility to the program. In addition, we computed an index of household assets used it as a proxy for permanent income. Our index was created using a factor analysis technique based on a list of 27 household appliances reported in the ECV/97.[12] In this case, a lower negative number, a larger value, implies higher assets.

Finally, a vector of community and local supply-side variables that affect the net benefits of the program, as well as the government's ability and willingness to enroll low-income families, was constructed and includes: the availability of a health center in the community (Red de Solidaridad Social), the existence of government-sponsored grassroots organizations in the community (*Red de Solaridad Social*), voter turnout in the 1994 municipal elections, the county's Living Standards Index (*Indice de Condiciones de Vida*), and the percentage of the state's population affiliated to the subsidized health system. All these variables were tested in the IV analysis as potential exclusion restrictions.

In addition, we add a proxy for the county's per capita health budget,[13] a dummy variables to indicate the region of the country and whether the community is rural or urban. Notice that these variables were used in the analysis as control variables in the participation and medical care use equations.

We hypothesize that the existence of a health center affiliated to the subsidy system in a community would increase the individual expected benefits of participating in the program. These types of providers would differ from providers in the safety net because they would give priority to patients within the SUBS system. Higher availability of these providers would improve access to care for SUBS enrollees.

Moreover, the level of affiliation in the SUBS program will capture the costs of searching for information regarding the program. Basically, higher aggregate level of participation in the SUBS program would imply that the information about the system will be more disseminated among potential beneficiaries. Individual political participation may indicate that households in areas with higher political turn out have better knowledge regarding the gains and costs of social programs in their communities. Both variables were treated as potential exclusion restrictions in the instrumental variable analysis.

On the other hand, whether or not a community is rural would affect the individual's decision to seek medical care as well as the individual's probability of being selected in the SUBS since municipal authorities are required to put an emphasis on enrolling individuals in rural communities according the general guidelines of the SISBEN program. In addition, we used a Life Standard Index as a proxy for the level of poverty at the municipal level. One may argue that communities with higher level of poverty will have a higher number of potential beneficiaries of the SUBS program.

Table 2 shows the means and standard deviation for selected variables used in the analysis. One should notice that some of the community and regional variables where used in the PSM analysis to form better matching between the treated and the control while the same variable were tested as potential exclusion restrictions in the IV analysis.

5. Results

Before discussing whether the subsidized health program has an impact on medical care use among participants, we first review the specification of the propensity score function, and its estimated values.

5.1. Individual Participation in the SUBS Program

The estimates of individual participation using unconstrained logit models are useful for two reasons. First, it gives some insight regarding the observable variables that should be included in the balancing function. Second, it provides a better understanding of participation of the low-income uninsured people in the program. Yet, using these data, we could not sort out if low-income people are opting out of the program, or if the county's authorities are not selecting them. Furthermore, based on this analysis, we could not interpret the real structural effect of the variable of interest on individual participation. For instance, a positive effect on the education variable may indicate that individual with higher education have lower costs to enroll in the program; however, it may also indicate that individuals with higher education perceive higher health benefits of participating in the program because the have higher preference toward medical care.

Table 3 shows the logit estimation of individual participation in the SUBS system. By all age groups, the result indicate: uninsured individuals living in households with lower levels of wealth, without telephones or bathrooms, living in households with fewer rooms, less education, and with bigger family sizes are more likely to participate in the program. However, since we do not have a direct measure of an individual permanent income, it is pertinent to evaluate other individual's characteristics before concluding that the program is really targeting the poorest members of the society. In particular, once one controls for household wealth, individuals who are married and employed are more likely to participate. Thus, both results suggest that among the uninsured, the individuals who are married and employed are more likely to be enrolled in the program.

Table 2. Descriptive statistics for selected variables.

		Treatment		Control	
Variable	Definition	Mean	Std dev	Mean	Std dev
	Outcome variables				
Preventive	= 1 if preventive care visit in the last 12 months, 0 otherwise	0.307	0.462	0.315	0.465
Hospital	= 1 if hospitalization in the last 12 months, 0 otherwise	0.065	0.247	0.053	0.223
Outpatient Visit	= 1 if outpatient visit in the last 30 days, 0 otherwise	0.196	0.397	0.165	0.371
	Individual characteristics				
Male	= 1 if male, 0 otherwise	0.491	0.499	0.513	0.499
Age	Individual's age	26.7	21.3	25.7	19.8
Married	= 1 if married, 0 otherwise	0.181	0.384	0.149	0.356
	Self-reported health status				
Very Good*	= 1 if self-reported health status is very good, 0 otherwise	0.074	0.262	0.092	0.289
Good	= 1 if self-reported health status is good, 0 otherwise	0.518	0.499	0.557	0.497
Regular	= 1 if self-reported health status is regular, 0 otherwise	0.347	0.476	0.306	0.461
Poor	= 1 if self-reported health status is poor, 0 otherwise	0.061	0.238	0.045	0.207
Chronic Condition	= 1 if individual reports a chronic condition, 0 otherwise	0.119	0.324	0.095	0.294
	Schooling				
None*	= 1 individual has no schooling, 0 otherwise	0.165	0.371	0.118	0.322
Elementary	= 1 individual has some elementary education, 0 otherwise	0.389	0.487	0.318	0.466
High School	= 1 individual has some high school education, 0 otherwise	0.072	0.258	0.169	0.376
College	= 1 individual has some College education, 0 otherwise	0.002	0.052	0.014	0.117
Employment	= 1 if the individual is employed, 0 otherwise	0.299	0.458	0.327	0.469
Self-employed	= 1 if individual is self-employed, 0 otherwise	0.192	0.394	0.218	0.413
	Household characteristics				
Household size	Number of people in the household	5.901	2.526	5.547	2.544
Worker ratio	% of household members that are employed	0.293	0.198	0.285	0.217
Assets	Index of household assets (from principal components analysis)	−1.021	0.579	−0.625	1.737
Electricity	= 1 if the house has electricity, 0 otherwise	0.821	0.384	0.875	0.331
Bathroom	= 1 if water-sealed toilet, 0 otherwise	0.221	0.414	0.453	0.498
Rooms	Number of rooms in the household	2.894	1.289	3.245	1.537
Floor_1*	= 1 if dirt floor, 0 otherwise	0.296	0.456	0.178	0.383

(*continued on next page*.)

Table 2. (*continued*).

Variable	Definition	Treatment Mean	Treatment Std dev	Control Mean	Control Std dev
Floor_2	= 1 if floor of unfinished wood or cement, 0 otherwise	0.616	0.486	0.553	0.497
Floor_3	= 1 if floor finished wood, tile, brick, or carpet, 0 otherwise	0.088	0.284	0.269	0.443
	Local characteristics				
Geographical Region					
Atlantica*	= 1 if living in Atlantica or San Andres, 0 otherwise	0.178	0.383	0.281	0.449
Oriental	= 1 if living in Oriental, 0 otherwise	0.288	0.453	0.121	0.325
Pacifica	= 1 if living in Pacifica, 0 otherwise	0.175	0.383	0.181	0.385
Central	= 1 if living in Central, 0 otherwise	0.151	0.357	0.167	0.373
Antioquia	= 1 if living in Antioquia, 0 otherwise	0.173	0.378	0.158	0.365
Bogota	= 1 if living in Bogota, 0 otherwise	0.018	0.135	0.065	0.246
Orinoquia	= 1 if living in Orinoquia, 0 otherwise	0.017	0.128	0.029	0.168
Rural	= 1 if living in an rural area, 0 otherwise	0.618	0.485	0.502	0.499
Health Expenditure	Municipal per capita health expenditure (1993 and 1997 average)	4.954	4.057	4.697	3.760
Health center^	= 1 if the community has a health center, 0 otherwise	0.522	0.499	0.624	0.484
Grassroots^	= 1 if community is part of *Red de Solaridad Social*, 0 otherwise	0.151	0.357	0.101	0.301
ICV 93^	Municipal's Living Standards Index for 1993	58.1	13.0	64.3	14.1
Election 94^	Voter turn out in 1994 Municipal elections	0.493	0.085	0.483	0.101
SUBS affiliation^	% of State's population affiliated to subsidized regime in 1996	0.187	0.058	0.156	0.051
Sample Size		5,559		16,732	

Notes: Individual variables not reported: mother's education, head of household's age and gender, head's health status, head's employment status, head's sector of employment. Housing variables not reported telephone connection, wall material.

*Category excluded in regression.

^ Potential exclusion restrictions.

As shown in Table 3, uninsured children between 0 to 4 years olds who are living in households with lower level of assets are more likely to be enrolled in the program. However, this result is not statistically significant. The evidence is somewhat mixed when one looks at other proxies of socio-economic status at the household level. In particular, children in households without bathrooms, and with fewer rooms are more likely to be in the program. Interestingly, controlling for socio-economic status, children with poor self-reported health status or with a chronic condition are more likely to participate in the program. Both results are statistically significant at $p < 0.05$. Both results are consistent with the idea of lower costs for the head of the household of searching for information to enroll their dependents in the program. Areas with higher amount of health transfer are more likely to enroll a child in the program.

Table 3. Logit estimations of program participation (treatment = 1). Estimated coefficients for selected variables.

Variables	Age 0–4 (a) Coefficient		Age 5–15 (b) Coefficient		Males Age 16–59 (c) Coefficient		Females Age 16–59 (d) Coefficient		Age 60–98 (e) Coefficient	
Constant	−1.184	(0.908)	−1.534	(0.589)***	−1.321	(0.583)**	−1.023	(0.554)*	−0.136	(1.644)
Individual characteristics										
Male	−0.154	(0.104)	0.078	(0.065)	NA	NA	NA	NA	−0.311	(0.174)*
Age	0.224	(0.038)***	−0.003	(0.011)	0.004	(0.003)	0.011	(0.003)***	−0.009	(0.001)
Married	NA	NA	NA	NA	0.223	(0.087)***	−0.003	(0.077)	0.134	(0.161)
Self-Reported health status										
Good	0.009	(0.168)	−0.068	(0.113)	0.311	(0.141)**	−0.114	(0.149)	0.007	(0.660)
Regular	−0.026	(0.196)	−0.182	(0.128)	0.087	(0.147)	−0.283	(0.153)*	−0.021	(0.658)
Poor	−1.067	(0.473)**	−0.257	(0.254)	0.108	(0.211)	−0.369	(0.193)*	−0.189	(0.677)
Chronic	0.809	(0.279)***	0.341	(0.166)**	0.532	(0.117)***	0.108	(0.098)	0.303	(0.162)*
Schooling										
Elementary	NA	NA	−0.051	(0.111)	−0.007	(0.085)	0.019	(0.081)	0.038	(0.154)
High School	NA	NA	−0.373	(0.359)	−0.403	(0.119)***	−0.279	(0.108)**	−0.751	(0.652)
College	NA	NA	NA	NA	−0.259	(0.374)	−0.755	(0.446)*	(f)	(f)
Employment	NA	NA	NA	NA	0.109	(0.109)	0.105	(0.101)	0.149	(0.317)
Self-employed	NA	NA	NA	NA	0.064	(0.076)	0.002	(0.110)	0.217	(0.284)
Household characteristics										
Household size	0.042	(0.024)*	0.053	(0.015)***	0.072	(0.014)***	0.059	(0.013)***	0.124	(0.031)***
Worker ratio	0.232	(0.418)	−0.303	(0.236)*	−0.436	(0.172)**	−0.293	(0.177)*	−0.526	(0.331)
Assets	−0.003	(0.079)	−0.142	(0.054)***	−0.129	(0.049)***	−0.135	(0.039)***	−0.127	(0.224)
Electricity	0.369	(0.156)**	0.249	(0.099)**	0.364	(0.105)***	0.196	(0.107)*	0.387	(0.223)
Bathroom	−0.634	(0.166)***	−0.535	(0.100)***	−0.652	(0.101)***	−0.649	(0.092)***	−0.018	(0.219)
Rooms	−0.106	(0.045)**	−0.131	(0.027)***	−0.145	(0.027)***	−0.149	(0.026)***	−0.251	(0.065)***

(*Continued on next page*.)

Table 3. *(Continued).*

Variables	Age 0–4 (a) Coefficient		Age 5–15 (b) Coefficient		Males Age 16–59 (c) Coefficient		Females Age 16–59 (d) Coefficient		Age 60–98 (e) Coefficient	
Community characteristics										
Rural	−0.301	(0.141)**	−0.283	(0.086)***	−0.418	(0.089)***	−0.234	(0.085)***	0.321	(0.195)*
Health expenditures	0.0391	(0.017)**	0.023	(0.011)**	0.015	(0.001)	0.031	(0.011)***	0.001	(0.002)
Health center	−0.206	(0.129)	−0.285	(0.079)***	−0.209	(0.081)**	−0.195	(0.076)***	0.123	(0.179)
Grassroots	0.351	(0.159)**	0.705	(0.102)***	0.431	(0.101)	0.438	(0.101)***	−0.062	(0.244)
ICV 93	−0.007	(0.005)	0.008	(0.003)*	−0.008	(0.003)**	−0.007	(0.003)**	−0.006	(0.008)
Election 94	−4.651	(1.412)***	−3.058	(0.876)***	−2.191	(0.877)**	−1.909	(0.839)**	−2.201	(2.111)
SUBS Affiliation	12.116	(1.487)***	9.651	(0.932)***	8.076	(0.929)***	7.837	(0.927)***	6.883	(2.103)***
LR Chi square	358.83***		822.6***		988.64***		969.17***		243.89***	
Pseudo R_square	0.136		0.127		0.151		0.141		0.173	
N	2,704		6,003		6,269		6,171		1,130	

Notes: ***Significant at $p < 0.01$, **Significant at $p < 0.05$, *Significant at $p < 0.1$.
(a) and (b) The variables included but not reported are; head of the household age, gender, employment status, type of employment, sector of employment and self-reported health, individual's mother education, telephone, floor and wall material, and 6 regional dummies.
(c), (d) and (e) The variables included but not reported are:sector of employment, telephone, floor and wall material, and 6 regional dummies.

For both males and females 16 to 59 age groups, there is a strong selection in favor of poorer uninsured individuals in terms of assets, and other household characteristics. Yet, the evidence indicates that holding other socio-economic characteristics constant, both males and females uninsured individuals with chronic health conditions are more likely to participate in the program. In addition, both females and males with some high school education are less likely to be selected into the program than individuals with no schooling. Surprisingly, individuals with more education are less likely to participate despite that they may face lower costs to search for information about the program and to enroll in the program. This may be consistent with the fact that people with higher level of education (i.e. higher permanent income) will be more efficient producers of health, and therefore, may demand less medical care. Married males are more likely than non-married males to be enrolled in the program; yet, married females are less likely to participate.

A similar pattern of results is reported for individuals older than 60 years (see Table 3). Note that, for the elderly population, employment status and level of schooling are not significant variables in the individual's decision to be in the program. Nevertheless, married individuals in this age group are more likely to be eligible in the program than non-married individuals; although this result is not statistically significant. Hence, it may suggest that the poorest elderly individuals are not enrolled in the program.

In sum, the results indicate that the individual participation depends on both individual characteristics to look or search for the program as well as on the government's capacity to offer the program in certain communities. In particular, in all estimation, the community's characteristics are a significant factor to explain program participation. These results suggest that one should consider both types of variables when specifying the propensity score function.

5.2. The Estimated PSM Function

Table 4 reports the main characteristics of the estimated propensity score functions by age group. The mean for each estimated function varies from 0.217 to 0.331, with little variability within each age group. Although not reported in the table,[14] the mean of the PSM estimated function for urban and rural also show low variability and it ranges from 0.292 and 0.142. The area of common support (similar propensity scores) between treated and control groups is high enough, ranging from 54.6% (for males individuals age 16–59 years) to 85.5% (for those less than 4 years old). Given the size of the available data, we were able to compute at least 5 blocks by age group; for the rural and urban data, the number of blocks were 10 and 9 respectively.

We were able to calibrate, for all age groups, and for the rural and urban groups, the function $z(x)$ until the balancing property was satisfied at significance level of $p < 0.005$. In each logit model, we managed to keep variables that capture the main factors (e.g., individual, household, and community characteristics) that affect the individual's probability of participating in the program (see Table 4 for a complete list of the variables).

In sum, the results of the propensity score functions for all age groups as well as for the rural and urban data suggest a high level of overlapping between the treated and controls

Table 4. Description of the estimated propensity score matching function $z(X)$ by age group.

	Age 0–4 (a)		Age 5–15 (b)		Age 16–59 Males (c)		Age 16–59 Females (d)		Age 60–98(e)	
Region common support	[0.032, 0.887]		[0.043, 0.644]		[0.011, 0.557]		[0.0177, 0.824]		[0.032, 0.826]	
Mean	0.266		0.307		0.217		0.248		0.331	
Std. Deviation	0.172		0.123		0.117		0.161		0.194	
Significance of balancing property	0.005		0.005		0.005		0.005		0.005	
Number of blocks	7		8		6		8		5	
	Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Control
Observations per-block	342	29	135	7	1,187	65	1,132	89	298	39
	550	102	307	38	695	103	1,290	222	240	102
	332	92	602	157	691	169	1,069	340	148	135
	206	110	504	119	1,300	421	565	263	31	78
	196	188	500	215	757	403	304	253	1	3
	34	71	828	430	256	195	131	201		
	2	10	733	630			64	130		
			5	12			1	7		

Notes: (a) The Propensity Score Function included the following variables: age, male, health status, chronic index of assets, electricity, telephone, bathroom, room, floor, wall, family size, worker ratio head of the household age and gender, head's health, head's employment status, head's sector of employment 6 regional dummies, urban, health center, Grassroots, election 94, ICV 93, SUBS afiliation.
(b) The Propensity Score Function included the following variables: age, male, health status, chronic index of assets, floor, family size, head of the household employment and head's sector of employment region2, region3, urban, health center, health expenditures, ICV 93.
(c) The Propensity Score Function included the following variables: age, health status index of assets, electricity, floor, family size, worker ratio employment, education, married, Orinoquia, urban, health center, Grassroots, ICV 93.
(d) The Propensity Score Function included the following variables: age, health status, chronic index of assets, electricity, telephone, rooms, floor, wall, family size, worker ratio employment, type and sector of employment, education, married, 6 regional dummies, urban health center, health expenditures, Grassroots, election 94, ICV 93, SUBS afiliation.
(e) The Propensity Score Function included the following variables: age, male, health status, chronic index of assets, electricity, telephone, rooms, floor, wall, family size, worker ratio employment, type of employment, education, married, 6 regional dummies, urban health center, health expenditures, Grassroots, election 94, ICV 93, SUBS afiliation.

in terms of pre-treatment observable covariates. Furthermore, the quality of the data allows us to match a high number of comparable observations within groups. Hence, applying PSM in this situation can offer a reliable initial estimate of expected treatment effect on the treated.

5.3. Does SUBS Increase Medical Care Use?

Table 5 reports our estimates for the treatment effect on the treated (ETT) based on the propensity score matching method. The bootstrapped standard errors are reported in parentheses. We constructed lower and upper bound statistically significant estimates of the program effect on the treated. The non-significant results were excluded when constructing these intervals. Finally, for comparison, the last column repeats the expected treatment effect (ET) already reported in Table 1.

Overall, our estimates support the hypothesis that the subsidized health insurance program increases medical care use among participants. In only three instances (outpatient care in 0–4 years old, hospitalization in 5–15, and preventive care in 60–98), there is no significant difference in ETT between treatment and control groups, regardless of the matching method used, which is consistent with the unconditional means for medical care use (i.e., the ET). For all other age ranges and measures of medical care use, there is at least one matching method that suggests a significant difference in ETT between the treatment and control groups.

Looking at the estimates in more detail, it is clear that there is considerable variability in ETT across matching methods. Nevertheless, the direction of the program's impact is positive for almost all cases, regardless of the matching method used. Furthermore, if we use a relatively loose standard ($p < 0.10$) and focus our attention on the range of estimated ETT that are statistically significant, we find that estimates for the average effect (ET) consistently underestimate the impact of the program.

Comparing the impact across age groups, we find that, whether or not the ETT is statistically significant, those over 59 years of age report a larger gain than women 16–59 years old and, in turn, these women report a larger gain than their male counterparts. This is true for all three measures of medical care use, regardless of the matching methods. For younger individuals, those under 16 years of age, the picture is quite mixed and no generalization can be made. Nevertheless, if we put aside the results from the radius method, then it is clear that the greatest gain in preventive care is obtained by the youngest group (0–4 years old), followed by the 5–15 years old.

Ideally, one would like to split our sample further and, for each age/gender group, conduct separate estimations of the ETT in urban and rural areas. In our case, our sample size is not large enough to perform such a detailed analysis.[15] Instead, we collapse all age groups and look for any impact differences between urban and rural areas. Table 6 reports our estimates for the ETT, for each measure of medical care use and for each of the four matching methods. We find that, in virtually all cases in both rural and urban areas, the ETT is positive and statistically significant and the the ET tends to underestimate the impact of the program, which is consistent with our previous findings by age group.

Individuals in urban areas obtain greater gains in all types of medical care uses. For instance, participants in urban areas tend to use more preventive care than non-participants. These differences in use range from 4.8% to 5.9%. Comparatively, the gains in preventive care use for participants in rural areas range from 2.9% to 4.8%. For outpatient care, the increase in urban areas is approximately 6% while in rural areas this value is smaller, around 4%. Finally, regarding hospital use the gains for participants in urban areas compared to

Table 5. Propensity score matching results—estimated treatment effect on the treated (ETT) in medical care use.

Age/Service	Matching method: Stratification		Nearest neighbor		Radius		Kernel		Lower and Upper Bounds (sig. at $p < 0.1$)	Difference in ET (from Table 1)
Age 0–4										
Preventive	0.074	(0.023) ***	0.074	(0.039)	0.023	(0.077)	0.070	(0.024)***	[0.070, 0.074]	0.023
Outpatient	0.017	(0.018)	0.012	(0.032)	−0.010	(0.065)	0.014	(0.019)	–	−0.006
Hospital	0.036	(0.011) ***	0.025	(0.016)	0.019	(0.033)	0.034	(0.011) ***	[0.034, 0.036]	0.022**
Age 5–15										
Preventive	0.068	(0.012) ***	0.063	(0.022) ***	0.056	(0.026) **	0.062	(0.015) ***	[0.056, 0.068]	0.012
Outpatient	0.031	(0.011) **	0.029	(0.016)*	0.028	(0.011) **	0.030	(0.009) ***	[0.028, 0.031]	0.020 **
Hospital	−0.001	(0.005)	0.000	(0.008)	−0.007	(0.009)	0.000	(0.005)	–	0.000
Age 16–59 Males										
Preventive	0.025	(0.014)*	0.018	(0.021)	0.011	(0.022)	0.021	(0.015)	0.025	−0.018
Outpatient	0.047	(0.012) ***	0.049	(0.018) **	0.057	(0.019)***	0.046	(0.011) ***	[0.046, 0.057]	0.037***
Hospital	0.014	(0.007) **	0.018	(0.011)*	0.007	(0.010)	0.015	(0.007) **	[0.014, 0.018]	0.011*
Age 16–59 Females										
Preventive	0.032	(0.015) **	0.037	(0.023)	0.029	(0.028)	0.025	(0.014)*	[0.025, 0.032]	−0.027*
Outpatient	0.077	(0.014) ***	0.092	(0.023) ***	0.089	(0.024) ***	0.079	(0.013) ***	[0.077, 0.092]	0.053***
Hospital	0.031	(0.010) ***	0.023	(0.014)*	0.027	(0.020)	0.027	(0.010) **	[0.023, 0.031]	0.023**
Age 60–98										
Preventive	0.041	(0.033)	0.048	(0.046)	0.110	(0.163)	0.041	(0.035)	–	−0.035
Outpatient	0.096	(0.031)***	0.118	(0.045)**	0.189	(0.147)	0.089	(0.029)***	[0.089, 0.118]	0.035
Hospital	0.031	(0.022)	0.070	(0.031) **	0.077	(0.092)	0.036	(0.022)*	[0.036, 0.070]	0.026*

Notes: $^{***}t$ value significant at $p < 0.01$, $^{**}t$ value significant at $p < 0.05$, $^{*}t$ value significant at $p < 0.1$.
Bootstrapped Standard Errors in parenthesis. The common support condition was imposed in all estimations.
Nearest Neigbor method refers to the random draw version.
In the Radius method, the size of the radius is 0.0001 Kernel method refers to the Gaussian Kernel

Table 6. Propensity score matching results—Estimated treatment effect on the treated (ETT) in medical care use. (Rural and Urban)

	Matching method									
Age/Service	Stratification		Nearest neighbor		Radius		Kernel		Lower and upper bounds (sig. at $p < 0.1$)	Difference in ET (from Table 1)
Rural										
Preventive	0.040	(0.010)***	0.048	(0.013)***	0.029	(0.013)**	0.037	(0.009) ***	[0.029, 0.048]	0.003
Outpatient	0.039	(0.008)***	0.041	(0.012)***	0.036	(0.010)***	0.039	(0.008)***	[0.036,0.041]	0.031***
Hospital	0.012	(0.005)**	0.018	(0.007)**	0.014	(0.007)	0.012	(0.005)**	[0.012,0.018]	0.013***
Urban										
Preventive	0.057	(0.011)***	0.057	(0.021)**	0.059	(0.018)***	0.048	(0.011)***	[0.048,0.059]	0.002
Outpatient	0.065	(0.011)***	0.063	(0.017)***	0.062	(0.016)***	0.063	(0.010)***	[0.062,0.065]	0.046***
Hospital	0.018	(0.006)**	0.011	(0.009)	0.02	(0.009)**	0.018	(0.007)**	[0.018, 0.020]	0.017***

Notes: ***t value significant at $p < 0.01$, **t value significant at $p < 0.05$, *t value significant at $p < 0.1$.
Bootstrapped Standard Errors in parenthesis. The common support condition was imposed in all estimations.
Nearest Neigbor method refers to the random draw version.
In the Radius method, the size of the radius is 0.0001.
Kernel method refers to the Gaussian Kernel.

non-participants go up to 2% while participants in rural areas have lower benefits compare to non-participants (1.8%).

Although urban populations obtain a greater gain in ETT than rural populations, this is probably the result of higher availability of health centers and other service providers in urban areas compare to rural areas; rather than the existence of shortcomings in the subsidy scheme per se in rural areas could explain this results.

On a more technical note, it is hard to tell which matching method, if any, performs best. For example, very few ETT in Table 5 that were estimated with the radius method turned out to be statistically significant. Such finding, rather than negating the results from the other methods, probably indicates that the length of the radius used (0.0001) is just too restrictive, resulting in too many observations being excluded from the common support area. In fact, if we turn to Table 6, where all age groups are condensed, all the ETT estimated with the radius method are statistically significant.

In sum, using the propensity score method, we find a greater impact of the program than what is suggested by simple differences in unconditional means. Since our data allows for a wide overlap in a relevant set observable covariates between the treated and controls, these estimated ETT give a more accurate approximation to the real program impact. Performing sensitivity analysis with different matching methods gives us an idea of the lower and upper bound of such an impact.

5.4. The IV Estimates of Program Participation on Medical Care Use

Tables 7 reports the results for the standard econometric analysis of the effect of program participation on medical care use where IV's are implemented to correct for the endogeneity of program participation. For the purpose of comparing the relevance of unobserved factors, the first column in the table shows the OLS estimates of the program effect on medical care use. The last two columns show the exclusion restrictions used in each regression, and the statistical test for the relevance of these variables as instruments.

Over-all, the analysis indicates that it is relevant to control for the endogeneity of program participation. In most instances, the F-test is significant at least at 10% level suggesting that IV estimates are consistent estimators. Furthermore, in most cases 2SLS estimates suggest that program participation seems to be relevant and with a significant positive effect in the individual's decision to use medical care. Finally, in most cases the direction of the results is consistent with the findings from the PSM analysis.

For instance, children 0–4 years old who are enrolled in the program tend to use more preventive and outpatient care than non-participants. According to our IV results, there is not a significant difference between groups in terms of the use of hospital services. For children 5–15 years old, the difference persists in terms of preventive care and hospital services.

According to the results, the program tends to benefit more female participants than non-participants between 16–59 years of age, in reference to preventive and outpatient care. Lower benefits are found for males in the program compared with non-participants in the same age group. Surprisingly, for both males and females, there are no differences between participants and non-participants in terms of utilization of hospital services. Finally, as

Table 7. OLS and instrumental variable estimates of program participation on medical care use.

Age/Service	OLS Estimates	IV Estimates	Instrumental Variables (exclusion restrictions)	F-value significant test for endogeneity (b)
Age 0–4				
Preventive	0.056 (0.019)***	0.304 (0.155)**	SUBS affiliation	2.85 *
Outpatient	0.029 (0.013)**	0.241 (0.113)**	SUBS affiliation	3.89 **
Hospital	0.023 (0.009)**	0.057 (0.075)	SUBS affiliation	0.22
Age 5–15				
Preventive	0.062 (0.012)***	0.774 (0.276)***	SUBS affiliation, ICV93, Election 94	9.87***
Outpatient	0.031 (0.008)***	0.828 (0.675)	SUBS affiliation, ICV93, Election 94	3.85**
Hospital	0.001 (0.004)	0.154 (0.088)*	SUBS affiliation, ICV93, Election 94	3.67*
Age 16–59 Males				
Preventive	0.042 (0.012)***	0.347 (0.161)**	ICV93, Election 94	4.22**
Outpatient	0.045 (0.009)***	−0.075 (0.116)	ICV93, Election 94	1.14
Hospital	0.021 (0.006)***	−0.132 (0.082)	ICV93, Election 94	3.73*
Age 16–59 Females				
Preventive	0.019 (0.013)	0.231 (0.113)**	ICV93, Election 94	5.16**
Outpatient	0.054 (0.012)***	0.035 (0.011)***	ICV93, Election 94	17.2***
Hospital	0.039 (0.008)***	−0.118 (0.071)	ICV93, Election 94	5.18 **
Age 60–98				
Preventive	0.031 (0.026)	−0.027 (0.481)	ICV93, SUBS affiliation	0.01
Outpatient	0.082 (0.025)***	0.385 (0.039)***	ICV93, SUBS affiliation	4.75**
Hospital	0.033 (0.016)**	0.025 (0.012)**	ICV93, SUBS affiliation	0.001

Notes: ****t* value significant at $p < 0.01$, ***t* value significant at $p < 0.05$, **t* value significant at $p < 0.1$.
(a) Standard errors are robust.
(b) F-value for the coefficient of the predicted program participation on the medical care use equation.
(c) For the age groups 0–4 and 5–15, the participation and medical care use equations include the following control variables : age, male, health status chronic, index of assets, electricity, telephone, bathroom, room, floor, wall, family size, worker ratio head of the household age and gender, head's health, head's employment status, head's sector of employment, 6 regional dummies, urban, expenditures.
(d) For males and females age 16–59, the participation and medical care use equations include the following control variables: age, health status chronic, index of assets, electricity, telephone, bathroom, room, floor, wall, family size, worker ratio employment, type and sector of employment, education, married, 6 regional dummies, urban, health expenditures.
(e) For age group 60 and older, the participation and medical care use equations include the following control variables: age, male, health status chronic, index of assets, electricity, telephone, bathroom, room, floor, wall, family size, worker ratio employment, type and sector of employment, education, married, 6 regional dummies, urban, health expenditures.

expected, an elderly individual who participate in the SUBS program consumes higher levels of outpatient care and hospital services than a non-participant elderly person. These results are largely consistent with the findings from our PSM analysis.

6. Conclusion

In this paper, our analyses suggest that the SUBS program in Colombia improves medical care use for children, women, and the elderly–groups that are of particular interest to policymakers. These results are consistent across methods (i.e., PSM and IV estimation). In general, our analysis suggests that although results from the methods are not directly comparable, the direction of the effect of program participation on medical care utilization is very similar for each medical care measures used in the analysis.

The PSM estimates suggest that the program has a larger positive effect than one may conclude based on a simple comparison of participants and non-participants. Furthermore, our results indicated that the screening of low-income families using decentralized systems and conducted by local authorities is successful in targeting the poor individuals in Colombia. Comparable systems for targeting the poor have been successfully used in other social programs in developing countries such as the Mexican Anti-poverty Program (PROGRESA). This empirical evidence suggests that there is potential to imitate similar demand-side government subsidy health insurance programs in other developing countries.

Further research should be conducted to evaluate the subsidy health insurance program in terms of the long-term gains in the health by low-income individuals. This paper assumes that enhancing medical care access would improve individual's health. Although that may be the case, there may be more cost-effective policy alternatives besides providing health insurance to low-income individuals.[16]

In terms of our methodology, PSM assumes the absence of unobservable variables that correlate with both the individual's decision to participate in the program and the outcome of interest. Despite this distributional assumption, this non-experimental methodology offers important advantages to the program evaluator. Furthermore, as this work suggests, comparing the results from the PSM analysis with IV's estimates could provide a more comprehensive and robust findings regarding the effect of subsidized health insurance program on medical care use.

PSM has been used with relative success in the evaluation of social programs in developing countries. For instance, using PSM Diaz and Handa (1999) evaluated the PROGRESA in terms of outcomes such as food expenditures, employment, and children's schooling. Their results indicate that simple matching effectively replicates the results from experimental studies (See Gertler, 2004 for an evaluation of PROGRESA using experimental data) as long the same survey instruments are used between the treated and controls. In this study, we met this condition.

The combination of the aforementioned results with our current finding suggests that policy makers could rely on this relatively inexpensive observational evaluation method to assess the impact of social programs in developing countries.

Acknowledgments

An earlier version of this paper benefited from comments by Gabriel Picone and Myron Fottler, as well as by participants in the International Health Economics Association (IHEA) meeting in 2003. We also would like to thank Juan Miguel Gallegos, Nelcy Paredes and Ursula Gideon for assistance in securing the household and community data. The authors take sole responsibility for any remaining errors. Antonio J. Trujillo would like to thank The University of Central Florida for the financial support provided for this work.

Notes

1. Bitran et al. (2003) estimate that by the end of 1997 about 20% of the population were in the subsidized regime but that by 2000 such percentage had declined to 16%. The overall picture is even worse if we take into account the country's economic recession during this period. As poverty indicators deteriorate and unemployment increases the enrollment in the contributory regime has declined, increasing the pressure over the subsidized regime. In fact, the percentage of population in the contributory regime declined from 34.3% in 1997 to 32.6% in 2000, along with an increase in the number of unemployed from 1.7 to 3 millions during the same period.
2. The SISBEN questionnaire has 62 questions. As a comparison, the 1997 Colombia Living Standards Survey has roughly 434 questions.
3. For a review of the issues common to most decentralized targeting programs see Conning and Kevane (2002).
4. Webpage of Colombia's *Departamento Nacional de Planeación*, as of September 30, 2003. http://www.dnp.gov.co/02_SEC/FOCAL/SISBEN.HTM
5. *Encuesta de Calidad de Vida*, conducted by Colombia's census bureau (DANE).
6. Panopoulus and Velez indicate that they tried to estimate a "Heckman type 'treatment effect' model", but failed probably due to small sample size (i.e., heads of household only).
7. Among the alternative approaches to solve the problem of missing data in non-experimental studies are: (i) instrumental variable estimators using cross sectional data; (ii) before and after estimators and (iii) difference-in-difference estimators. For a complete discussion of these alternative methods see Moffit (1991) and Blundell and Dias (2000).
8. For the implications and formal proof of the balancing property and the ignorable treatment assignment property see Rosenbaum and Rubin (1983, 1984).
9. Becker and Ichino (2002) developed this algorithm. We compared the results from this program with the results from the algorithm developed by Sianesi (2001). Comparable ETT results were similar.
10. *Caja* and EPS include various special insurance schemes covering teachers, policemen, the military, and employees of the national oil company *Ecopetrol*. Additional insurance programs are run by certain charity organizations such as *Canitas* and Plan *Revivir*.
11. The reader should keep in mind that this assumption about the control group overstates the estimated treatment effect since some fraction of the target population may have insurance. Nevertheless, this effect would be negligible. We thank one of the anonymous referees for this comment.
12. Estimation results are available from the authors upon request.
13. This variable includes the total health transfers from the Central Government to the municipalities. According to the current decentralization law, 25% of the total transfers from the central government must be allocated in health expenditures.
14. These results are available upon request from the author.
15. Note that the matching for each age group was conducted with the urban-rural dummy variable as part of the balancing property.
16. See Brook et al. (1983), for a similar argument in the U.S context.

References

Baker, J. (2000). *Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners*. The World Bank.

Becker, S. O. and A. Ichino. (2002). "Estimation of Average Treatment Effects Based on Propensity Scores." *Stata Journal* 2, 358–377.

Bitran & Asociados-Econometria S.A. -ESAP (2003). "Evaluacion y Restructuracion de los Procesos, Estrategias y Organismos Publicos y Privados encargados de la Afiliacion, Pago y Recaudo de Aportes al SGSS. Informe final.

Bloom, E., L. Mina and C. Olivo. (1998). *Eficiencia y Equidad en el Sector Salud Colombiano*. Informe final.

Blundell, R. and M. Costa Dias. (2000). "Evaluation Methods for Non-Experimental Data." *Fiscal Studies* 21(4), 427–468.

Brook, R., Ware, W. Rogers, E. Keeler, A. Davies, Cathy Donald, G. Golberg, K. Lohr, P. Masthay and J. Newhouse. (1983). "Does Free Care Improve Adults' Health?" *The New England Journal of Medicine* 309, 1426–1434.

Conning, J. and M. Kevane. (2002). "Community Based Targeting Mechanisms for Social Safety Nets: A Critical Review." *World Development* 30(3).

Davison, R. and J. G. MacKinnon. (1993). *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Dehejia, R. H. and S. Wahba. (1999). "Causal Effects in Non-experimental Studies: Reevaluating the Evaluation of Training Programs." *Journal of the American Statistical Association* 94, 1053–1062.

Departamento Nacional de Planeacion-DPN, Unidad de Desarrollo Social. Sistema de Seleccion de Beneficiarios para Programas Sociales (SISBEN).

Diaz, J. J. and S. Handa. (2004). "An Assessment of Propensity Score Matching as a Non-Experimental Impact Estimator: Evidence from a Mexican Poverty Program." Unpublished Manuscript.

Friedlander, D. and P. K. Robins. (1995). "Evaluating Program Evaluations: New Evidence on Commonly Used Non-experimental Methods." *American Economic Review* 85(2), 923–937.

Gertler, P. (2004). "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment." *American Economic Review* 94(2), 336–341.

Heckman, J. (1996). "Randomization as an Instrumental Variable Estimator." *Review of Economics and Statistics* 56, 336–341.

Heckman, J. (1990). "Varieties of Selection Bias." *American Economic Review* 80, 313–318.

Heckman, J. J. and V. J. Hotz (1989). "Choosing Among Alternative Non-experimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association* 84, 862–874.

Heckman J. J., H. Ichimura and P. E. Todd. (1997). "Matching As An Econometric Evaluation Estimator: Evidence From Evaluating a Job Training Program." *Review of Economics and Statistics* 64, 605–654.

Heckman, J. J. and J. A. Smith. (1995). "Assessing the Case for Social Experiments." *Journal of Economic Perspectives* 9, 85–110.

Ichimura, H. and C. Taber. (2001). "Propensity-Score Matching With Instrumental Variables." *AEA Papers and Proceedings* 91, 119–124.

Jack, W. (2000). "The Evolution of Health Insurance Institutions: Theory and Four Examples From Latin America." World Bank. Working Paper.

Jalan, J. and M. Ravallion. (2003). "Estimating the Benefit Incidence of an Antipoverty Program by Propensity Score Matching." *Journal of Business and Economics Statistics* 21, 19–31.

Jaramillo, I. (2001). "Evaluación de la descentralización en salud en Colombia." *Archivos de Economia*, # 170, Departamento Nacional de Planeación.

Joffe, M. M. and P. R. Rosenbaum. (1999). "Propensity Scores." *American Journal of Epidemiology* 150, 327–333.

LaLonde, R. J. (1986). "Evaluating the Econometric Evaluations of Training Programs With Experimental Data." *American Economic Review* 76, 604–620.

Londoño, B., I. Jaramillo and J. P. Uribe. (2001). "Descentralizacion y Reforma en los Servicios de Salud: El caso Colombiano." Working Paper. World Bank.

McKay, A. (2002). "Should the Survey Measure Total Household Income?" In M Grosh and P. Glewwe (eds), *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the LSMS*.

McGee, R. (1999). "Technical, Objective, Equitable and Uniform? A Critique of the Colombian System for the Selection of the Beneficiaries of Social Programs SISBEN." IDPM Discussion Papers, University of Sussex, Working Paper 57, 1–34.

Minkin, A. (2002). "Nonparametric Propensity Score Matching Method for Treatment Effect Estimation in Observational Studies." Working Paper. Department of Economics, University of Wisconsin.

Moffitt, R. (1983). "An Economic Model of Welfare Stigma." *The American Economic Review* 73, 1023–1035.

Moffitt, R. (1991). "Program Evaluation With Non-Experimental Data." *Evaluation Review* 15(3), 291–314.

Panopoulus, G. and C. E. Velez. (2002). "Subsidized Health Insurance, Proxy Means Testing and the Demand for Health Care among the Poor in Colombia." Working Paper, 214–241.

Pauly, Mark V. (1968). "The Economics of Moral Hazard: Comment." *American Economic Review* 3, 531–537.

Peñaloza, E. and E. Londoño. (2002). "Regimen Subsidiado de Salud en Colombia: Evaluacion y Propuestas de Ajuste. Fundacion Corona, Banco Interamericano de Desarrollo." CENDEX, Potificia Universidad Javeriana.

Remler, D. and S. Glied (2003). "What other Programs can Teach Us: Increasing Participation in Health Insurance Programs." *American Journal of Public Health* 93, 67–74.

Rosenbaum, P. R. and D. B. Rubin. (1983). "Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70, 41–55.

Rosenbaum, P. R. and D. B. Rubin. (1984). "Reducing Bias in Observational Studies Using Sub-classification on the Propensity Score." *Journal of the American Statistical Association* 79, 516–524.

Rosenbaum, P. R. and D. B. Rubin. (1985). "'Bias Due to Incomplete Matching." *Biometrics* 41, 103–116.

Sianesi, B. (2001). "An Evaluation of the Active Labor Market Programmes in Sweden Office of Labor Market Policy Evaluation." Institute for Labor Market Policy Evaluation. Working Paper.

Sianesi, B. (2001). "PSMATCH: Stata Module to Perform Various Types of Propensity Score Matching." http://ideaas.repc.org/p/boc/bocode/s418602.html.

Smith, J. A. and P. E. Todd. (2001). "Reconciling Conflicting Evidence on the Performance of Propensity-Score Matching Methods." *American Economic Review* 91, 112–124.

Trujillo, A. (2003). "Medical Care Use and Selection in a Social Health Insurance with an Equalization Fund: Evidence from Colombia." *Health Economics* 12, 231–246.

Velez, C. E., E. Castano and R. Deutsch. (1999). "An Economic Interpretation of Targeting Systems for Social Programs: The Case of Colombia's SISBEN." World Bank. Working Papers.

Velez, C. E. and V. Foster. (2000). "Colombia Poverty Study. The Impact of Public Social Expenditure." Latin American Caribbean Region, World Bank, Washington D.C.