

SMALL-AREA VARIATION IN HEALTH CARE AFFECTING THE
CHOICE OF CESAREAN DELIVERY: THE CASE OF A COLOMBIAN
HEALTH INSURER

Andrés I. Vecino-Ortiz
David Bardey
Ramón Castano-Yepes

SERIE DOCUMENTOS DE TRABAJO

No. 73

Octubre 2009

Small-Area Variation in Health Care Affecting the Choice of Cesarean Delivery: The Case of a Colombian Health Insurer

Andrés I. Vecino-Ortiz¹, David Bardey², Ramón Castano-Yepes³

Abstract

In the midst of health care reform, Colombia has succeeded in increasing health insurance coverage and the quality of health care. In spite of this, efficiency continues to be a matter of concern, and small-area variations in health care are one of the plausible causes of such inefficiencies. In order to understand this issue, we use individual data of all births from a Contributory-Regimen insurer in Colombia. We perform two different specifications of a multilevel logistic regression model. Our results reveal that hospitals account for 20% of variation on the probability of performing cesarean sections. Geographic area only explains 1/3 of the variance attributable to the hospital. Furthermore, some variables from both demand and supply sides are found to be also relevant on the probability of undergoing cesarean sections. This paper contributes to previous research by using a hierarchical model and by defining hospitals as cluster. Moreover, we also include clinical and supply induced demand variables.

JEL Classification I10, I11, I18.

Keywords: *Small-area variation, health care, procedure of delivery, multilevel analysis.*

1 NYSPi-Columbia University Medical Center, New York, USA.

2 Universidad del Rosario, Bogotá D.C. (Colombia) and Toulouse School of Economics (France).

3 Caja Colombiana del Subsidio Familiar- Colsubsidio, Bogotá D.C., Colombia

Corresponding author: Andrés I. Vecino-Ortiz, email: andres.vecino@gmail.com.

The findings, recommendations, interpretations and conclusions expressed in this paper are those of the authors and not necessarily reflect the view of the Department of Economics of the Universidad del Rosario

1 Introduction

Developing countries are currently experiencing rising cesarean rates whereas countries with the best birth outcomes continue to have low cesarean section rates (Althabe *et al.*, 2006; Villar *et al.*, 2006). In 1985, the World Health Organization stated that cesarean rates should not be higher than 15%. Even though rule of thumb for almost 25 years old, it reflects that this is an alarming issue. Concerns about cesarean section rates arise from its potential medical complications and the rising costs the procedure represents for government and insurers. Cesarean section is a major surgical procedure and can cause medical complications including wound infection, respiratory distress, neonatal sepsis and death for mother and child (Iams and Chawla, 1984; MacDorman *et al.*, 2006; 2009; Tita *et al.*, 2009). Since diseases related to pregnancy and delivery are the main non-violent cause of death among women in reproductive age (Ministry of social Protection and Pan American Health Organization, 2007) and Colombia is ranked as the third highest country in Latin America in performing cesarean sections (30% cesarean sections on excess according to Belizán *et al.* in 1999), this should be considered as a high priority issue in this country.

As with other procedures in the health sector, the performance of the cesarean section has changing and controversial criteria, and most of them show a lack of rigorous efficacy and cost-effectiveness support. On the background of informational asymmetry between physicians and patients, this lack of evidence regarding the effectiveness and efficiency on the performance of cesarean sections leaves a grey zone, allowing health care agents (insurers, hospitals and physicians) to respond to the incentives they face.

Research on “*small-area variation in health care*” has become a relevant research topic since the Glover’s initial paper in 1938. Later, a cluster of research (performed mostly between 1970 and 1990) has assessed several procedures in health care in developed countries, with most of them taking place in the United States. In 1993, Phelps and Mooney brought out their seminal paper on regional variations, and became the referential framework to analyze variations in health care thereafter. “*Small-area variation in health care*” is usually defined as the systematic variability on the probability of performing a given procedure across “small areas”. The concept of “*small area*” may range from small geographical regions (such as counties and towns), to entire countries. However, because the definition of “*small area*” may not represent legal, financial or practice-style effects, each “*small area*” might not affect homogeneously the choice of a given procedure.

Concerns on small-area variation on health care arise from the potential welfare losses they produce. Despite the fact that some variability is acceptable between regions conditioned on disease prevalence, insurance coverage and availability of resources, systematic variability yields welfare losses in both users and non-users of a given service. More precisely, welfare losses in users arise from the production of non-optimal amounts of services for them with regards to more appropriate substitute services, even if no procedure is performed (Phelps and Mooney, 1993). In non-users, systematic variability generates welfare losses, because socially-inefficient allocation of resources creates social marginal costs that are greater than social marginal benefits (Phelps and Parente, 1990). The excessive or deficient provision of a given procedure generates welfare losses both directly and indirectly.

Since Colombia faced health reform 15 years ago, such sector restructure brought about considerable changes in the national health care market, increasing the coverage and supply of health services. However, through these years, as financing grows slower than the plan of benefits,⁴ the rationing of health care services according to efficiency criteria has become the first priority for insurers and government sector.

This paper attempts to contribute to “*small-area variation in health care*” research line in several ways. To our knowledge, this is the first paper using hierarchical econometric methods which takes hospitals uses as unit of analysis rather than geographic regions. The rationale behind this conceptual model is that hospitals’ policy and environment may exert a considerable influence on the choice of the physician in a determined procedure. In addition, earlier research has focused on developed countries; this paper can be used to fill the gap by studying this issue in developing countries. Finally, we control for supply induced demand and clinical data at the individual-level, which allows us to go through the effects of both variables on the choice of the delivery procedure.

Individual-level and hospital-level data on all births during 2007 from a Contributory Regimen health insurer (EPS-C) are used. As a result of the objectives previously stated, we do not aim to infer our results at the national level but only to use a highly reliable single-insurer database. The first level of analysis is comprised by births whereas the second one is identified by hospital. Two different specifications for the error term on a multilevel logistic regression model are performed. In the first specification, this paper finds that 20% of the variance on the performance of cesarean sections is attributable to differences between hospitals. In the second specification, we show that 1/3 of the variance attributable to the hospital is explained by Regional differences. The remaining 2/3 of this variance is due to the hospital clustering effect. We also found that wealthier and more educated mothers, as well as those attended in a private

⁴ POS, Plan Obligatorio de Salud.

facility, are more likely to undergo a cesarean section. In addition, we found that physicians' fees are positively correlated with the performance of cesarean sections, confirming a supply induced demand effect.

The paper is organized as follows. Section 2 presents a brief literature review of research on "*small-area variations on health care*" and defines an approach to this issue through the seminal work of Phelps and Mooney (1993). Section 3 is devoted to the presentation of the data and includes descriptive statistics. Section 4 presents the empirical strategy detailing both multilevel analysis specifications. Section 5 presents results from both multilevel specifications. In section 6, our results are discussed. Section 7 states our conclusions and policy recommendations.

2 Background

Small-area variation in health care is first described by Glover (1938) when he found that cities in the United Kingdom had different tonsillectomy rates with no evidence of different tonsillitis incidence. Through an analysis of the data, Glover concluded that some cities were performing tonsillectomies "*in excess of the essential minimum*". Later on, other influential papers on small-area variations using linear approaches and aggregate-level data have been published (Cromwell and Mitchell, 1986; Mc Pherson *et al.*, 1982). Using a panel data approach, Bech and Lauridsen (2007) find more than 50% of variability between municipalities. This paper constitutes an important contribution to this field, as they used longitudinal data.⁵

Phelps and Money (1993) provide a seminal paper explaining small area variations from a demand-supply side approach. In this analysis, both users and healthcare providers cause variations across geographic areas. This demand-supply side approach is followed in this paper hereafter.

Usually, approaches focusing on the demand side discuss two main aspects: patients' preferences and the demographic characteristics of population. The effect of consumers' preferences is highly relevant on the final choice of performing a given procedure since they might shape the decision either for or against a given procedure by persuading their physician. Even though patients' preferences are expected to have stronger effects under out-of-pocket payments (Bardey and Lesur, 2006), they also may play a role in their absence (Hsu *et al.*, 2008). Such preferences are mainly determined by cultural and "social

⁵ Nonetheless, panel data approach is not able to account for clustering effects.

inheritance” factors. Cultural factors are comprised mostly by the perception of cesarean section as a less painful, easier and a more convenient procedure than vaginal delivery (Fabri and Murta, 2002; Tschudin *et al.*, 2009). Empirical evidence for “social inheritance”, as called in Tollånes *et al.* (2008), shows that women are more likely to experience the same delivery procedure their mothers had. Although they do not discard the involvement of genetic features, they point out that “social inheritance” becomes a more plausible explanation for this phenomenon. Demographic characteristics are also important in determining the intensity of using health services, as some specific services might be affected by factors related to age, gender, or ethnicity. For instance, Internal Medicine Services are demanded more in elderly populations while more fertile populations demand more obstetric and pediatric services (Phelps and Money, 1993).

Small-area variations from the supply side are mainly determined by informational asymmetries that characterize health care markets (Phelps and Money, 1993). On the one hand, health care providers benefit from information asymmetries regarding quality, which is hardly observable by patients, making them less quality-sensitive and lowering competition. As health care providers make up the link between the patient and the insurer, they function as “double agents” (Blomqvist, 1991). Such asymmetries yield two main factors with effects on the variability of procedures; the incentives physicians face and their own preferences. Availability of technology and institutional guidelines also play a role in determining small-area variations (Phelps, 1993). Furthermore, physicians’ sensitivity to the remuneration generates supply induced demand behaviours. It has been largely studied that the remuneration scheme may affect the procedure that providers choose. Indeed, supply induced demand is the ability of the providers to deliver a quantity/quality of health care different than the one her patient would choose if she had the same information (Labelle *et al.*, 1994). This phenomenon involves a principal-agent problem where physician (agent) participates in the decision-making process and owns privileged information regarding the quality of the service and available technology (such as evidence-based medicine recommending or not a given procedure). Thus, the physician is able to take advantage of this informational asymmetry to recommend the most profitable choice for her. Roughly speaking, the physician maximizes her utility upon the chosen procedure, regardless of the patients’ utility (Pauly and Satterthwaite, 1981).

In addition to the incentives issue, physicians’ preferences are also important in determining the choice of procedures from the supply-side factor, even though there is much controversy regarding the evidence confirming such effect (Epstein and Nicholson, 2009; Grytten and Sørensen, 2003; Stano, 1991). Physicians’ preferences are often referred as “school effects” (Phelps and Mooney, 1993) or “practice styles” (Wennberg, 1985; Wennberg and Gittelsohn, 1973). These preferences are caused by differences in information regarding the available technology (which includes guidelines and evidence-based medicine) to which physicians are exposed during their training and professional practice. This

implies that two physicians may develop divergent criteria throughout their carrier regarding the effectiveness of a given intervention (Davis *et al.*, 2000; Newhouse and Phelps, 1976; Phelps and Mooney, 1993).

3 Data and descriptive statistics

3.1 Data

Despite our efforts in obtaining data at the national-level, we face a trade-off between heterogeneous, low quality and incomplete national-level data and homogeneous, high quality and reliable information from a single insurer. As a result, we chose to use a single insurer database⁶. This data comes from a private and for-profit insurance company operating at the national-level for the Contributory Regimen (EPS-C). Data coming from administrative and RIPS records (Individual Registry of Health Care Services) on all births happened during 2007 have been obtained (11,768). A sample comprising 98.5% of the original data was selected. As 16 births were delivered in low-complexity facilities where no infrastructure to perform cesarean sections was available (and therefore, no possibility of choosing the delivery procedure), these observations are excluded from the analysis. Five observations had missing information regarding the hospital in which the delivery happened. Births delivered at hospitals with less than 10 observations (142 observations) are also excluded to ensure the suitability of the model⁷. Moreover, during the period of analysis, some hospitals performed either only cesarean sections or only natural deliveries. As these might be hospitals agreeing contracts with the insurer only for specific procedures (for example, a high-complexity hospital might arrange a contract with the insurer to perform only cesarean sections referred from other hospitals), hospitals in which only one type of procedure is observed are also excluded from the analysis. Ultimately, 174 observations were excluded yielding a final sample of 11594 observations.

6 This choice is based on the aims of the paper, which are 1) to contribute to present research using a hierarchical econometric approach, which defines hospitals as clusters and also includes clinical and supply inducement data 2) to make Colombian policymakers aware that this is an actual issue that health sector is facing.

7. Snidjers and Bosker (1993) explain how a sample of at least ten individuals for each group is a conservative way to optimize the estimation efficiency in multilevel analysis.

3.1.1 Hospital variables

Variables at the hospital level are obtained in order to understand their effects on the choice of either cesarean section or natural delivery. Hospitals in Colombia are classified as low, medium and high complexity. Whereas low complexity facilities provide mostly ambulatory services, middle complexity hospitals provide in-patient and surgical services for general specialties (internal medicine, pediatrics, gynecology and general surgery). High complexity institutions (tertiary-care level) provide high-specialized services such as intensive-care units and subspecialties. As births delivered in low complexity facilities were excluded from the study, a dummy variable accounting for high complexity facilities is designed. *High complexity* is a variable taking a value of 1 if the delivery was attended at a tertiary-care hospital and 0 otherwise (medium-complexity hospital).

Following the previous literature (Grant, 2009; Gruber *et al*, 2009), as all physicians worked on a fee-for-service basis, a supply induced demand variable reporting the effect of physician remuneration on the choice of the delivery procedure is designed as follows:

$$DI_{ij} = (W_i - NatW_j)^{(Ces)} (W_i - CesW_j)^{(1-Ces)} \quad (1)$$

where W_i represents the actual fee the physician receives for carrying out the birth i . $CesW_j$ and $NatW_j$ are the average fees received by physicians working at hospital j for performing respectively a cesarean section or a natural delivery. Ces takes value of 1 if the delivery procedure was cesarean section and 0 otherwise. This variable accounts for the average difference between the actual and the alternative procedure's fee within each hospital. The rationale for this variable is to describe the incentive faced by the physician for the chosen procedure; for example, the value for performing a vaginal delivery is the differential fee between performing a vaginal delivery and a cesarean section at that hospital. As there is a considerable overlapping on the values of physician' fees for both cesarean sections and natural deliveries, there are no simultaneous identification of this variables with the dependent one.

Region of the hospital location is also categorized according to previous Colombian research on health issues (Castano and Zambrano, 2006; Vecino-Ortiz, 2008). Such regions were categorized as follows; Bogotá D.C, Central Region (Antioquia, Quindío, Risaralda, Caldas), Pacific Ocean Region (Valle del Cauca, Cauca and Nariño), Eastern Region and Caribbean Region (Bolívar, Magdalena and Atlántico). It is worth noticing that, as this is a single-insurer case, the patient distribution among departments does not correspond to the population distribution. *Public hospital* is a dummy variable taking a value of 1 if the

hospital is public and 0 if private. *Teaching hospital* is also a dummy variable that accounts for hospitals that conduct research, and have relationships with medical schools on a regular basis (Epstein and Nicholson, 2009; Grant, 2009; Jaramillo *et al.*, 2008).

3.1.2 Individual variables

A set of individual-level variables is used. *Mother age* is the mother's age at the time of birth. *Mother income* is the value in hundred thousand pesos of the reported income to the insurer during the last month. *Users per contract* is a variable accounting for the number of people covered under the family group. This variable is included as a proxy for family size which may play an effect as suggested in Tollånes *et al.* (2008). Educational level is controlled using three variables: *Primary education accomplished*, *Secondary education accomplished* and *Superior education level accomplished*. *Superior level education accomplished* includes technical and professional education. In addition, a variable for *unknown educational level* was included since in 4,922 observations education information is not provided. Mother's labor information is categorized as *Dependent worker*, *Self-employed worker* and *No working currently*.

In order to control for health factors, some clinical variables are also included. *Previous births* are the number of births a woman had before this delivery. *Type of admission* describes whether the admission was urgent or scheduled. *Female* takes values of 1 if the newborn is female and 0 if male. Even though data on prenatal care and other prevention activities and gestational age, were requested, these data were not available.

The last individual variable is *Cesarean section criterion*. The categorization was developed following Flamm and Quilligan (1995) and Ministry of Social Protection (2000). Since complications are listed in the database as ICD-10 codes (International Classification of Diseases, version 10), the imputation of whether that intervention requires cesarean section or not is based only on the ICD-10 diagnosis and on the common treatment upon such ICD-10 diagnosis. Two values are imputed for this variable, 1 if that complication requires cesarean section (according to previously discussed literature) and 0 if such complication either does not require cesarean section or is not related to the procedure.

3.2 Descriptive statistics

From this sample (11,594 births), 35.33% of births (4,096) are delivered by cesarean section. In Table 2, descriptive results for continuous variables are displayed. *Supply induced demand* achieved a mean of

-55,000 pesos⁸ with a standard deviation of 469,000 pesos (\$225.66). Mothers mean income is 534,000 pesos (\$256.93) with a standard deviation of 155,000 pesos (\$74.58). Mean of previous deliveries is 0.24. Average users per contract in this sample are 1.10.

Overall, as shown in Table 3, a high proportion of deliveries are attended at high-complexity, private and non-teaching hospitals, most of them at the Central region (Medellin). 80.9% of mothers were between 19 and 35 years old. Female/male proportions are as expected and almost 90% of births were admitted as urgent. 27.60% of mothers did not declare their level of education, whereas 42.45% of mothers did not complete primary education and 18.33% completed secondary education. Almost 90% of mothers were dependent workers and only 1% was not currently working. Regarding the cesarean section criteria, 356 cases (3.07% of sample) accomplished criteria for cesarean section.

Table 2. Descriptive Statistics of Continuous Independent Variables				
Variables	Summary statistics (n=11594)			
	Mean	Standard deviation	Minimum	Maximum
Hospital variables				
Supply induced demand*	-0.55	2.69	-14.78	14.78
Individual variables				
Mother income*	5.34	1.55	0.00	19.87
Previous deliveries	0.24	0.54	0.00	6.00
Users per contract	1.10	0.31	1.00	5.00
* Value in hundred of thousand pesos				

⁸ -\$26.43 taking into account the average exchange rate for 2007 of 2,078.35 pesos per U.S. dollar.

Table 3. Descriptive Statistics of Dummy Independent Variables

Variables	Proportions (n=11594)
	%
Hospital variables	
Hospital complexity	
High	91.75%
Intermediate	8.25%
Hospital location	
Bogota D.C	13.46%
Central region	72.35%
Pacific region	4.78%
Eastern region	1.92%
Caribbean region	7.49%
Type of hospital	
Public	10.76%
Private	89.24%
Teaching hospital	
Yes	13.37%
No	86.63%
Individual variables	
Mother age	
Less than 18 years old	10.04%
18 to 35 years old	80.90%
More than 36 years old	9.06%
Newborn sex	
Female	47.82%
Male	52.18%
Mother education	
Unknown mother education	27.60%
Mother education less than completed primary	42.45%
Mother completed primary education	7.01%
Mother completed secondary education	18.33%
Mother completed any kind of superior education	3.91%
Type of worker	
Dependent	89.05%
Self-employed	9.27%
No working currently	0.97%
Cesarean section criterion	
Accomplished	3.07%
Unaccomplished	96.93%

4 Empirical strategy

Multilevel logistic regression analysis is carried out since it provides estimates of the hospital-attributable variance on the probability of performing the delivery procedure. This method also shows the effects of a set of independent variables on such probability.

4.1 *Multilevel logistic regression model*

Multilevel logistic regression is the dichotomous dependent-variable variant of the Multilevel Regression Model. Multilevel regression models belong to the “Hierarchical Models Family” and are frequently used with data with complex variability patterns arising from clusters or nested data (Snijders and Bosker, 1999). The technical advantage of this methodology relies on the error term structure. Whereas linear or logistic regression models exhibit one error term for the whole equation, multilevel analysis generates one error term for each level, allowing the researcher to isolate the individual-level and group-level residual variance.

Overall, multilevel logistic regression model estimates the cluster-dependent probabilities’ distribution (variance) of the binary dependent variable. Multilevel logistic regression comprises two subtypes of models known as the “*empty model*” and the “*random-intercept model*”. The main difference between both models is that the first one does not take into account any explanatory variable, whereas the “*random-intercept model*” does it. It means that whereas in the “*empty model*”, the observed probability of success varies across the groups, in the “*random-intercept model*” such probability varies across individuals instead. As a result of the “*empty model*” parameters, the *Intraclass Correlation Coefficient* (ICC) may be calculated. The *Intraclass Correlation Coefficient* represents the fraction of variance on the binary-dependent variable attributable to the differences between groups (Rabe-Hesketh and Skrondal, 2008; Snijders and Bosker, 1999).

4.2 *Specification*

In the specifications this analysis performs, births are taken as units of observation (level-one units) and hospitals as groups of analysis (level-two units). As our interest is to estimate not only the hospital but also the region effect on the choice of the delivery procedure, region is initially considered as a potential third level in the model. However, a small number of regions where hospitals are located (five geographic regions) prevent us from an accurate analysis of region as a level, given that standard errors in

multilevel analysis are not well understood when the number of groups is small (Raundenbush, 1989; Snijders and Bosker, 1993). Nonetheless, we overcome this issue by using an alternative variance decomposition specification. Such specification (from now on, Model 2) takes the region as a random residual in order to isolate the residuals accounting for the region effect on the hospital variability from the remaining residuals (Rabe-Hesketh and Skrondal, 2008).

As a result, two different sets of models, differing on the residual structure specification, are carried out. In the following reduced models, Y_{ij} is the dummy of the dependent variable, taking values of 1 if a cesarean section is performed and 0 otherwise for the birth i in the hospital j . Whereas P_j represents the constant of the regression for the empty model, P_{ij} is a vector *that* characterizes both the constant and the set of independent variables with corresponding estimators for the random-intercept model.

Model 1

In the first specification, both empty and random-intercept models are performed with the residual R_{ij} accounting for the whole residual variance of the individual i within the group j .

- Empty:
$$Y_j = P_j + R_j \tag{2}$$

- Random-intercept:
$$Y_{ij} = P_{ij} + R_{ij} \tag{3}$$

Model 2

In the second specification, where variance decomposition is performed, μ_{ij} represents the residual accounting for the effect of the region and the ε_{ij} residual comprises the remaining effect of the individual i within the group j for $R_{ij} \approx \mu_{ij} + \varepsilon_{ij}$.

- Empty:
$$Y_j = P_j + R_j + \varepsilon_j \tag{4}$$

- Random-intercept
$$Y_{ij} = P_{ij} + R_{ij} + \varepsilon_{ij} \tag{5}$$

Variance inflation factor and conditional numbers techniques have been also performed to confirm the absence of multi-collinearity between variables.

5 Results

In this paper, a comparative assessment between two specifications on the probability of undergoing a cesarean section *versus* vaginal delivery is performed using multilevel analysis (including empty and random-intercept models).

Table 4. Variance Results and Intraclass Correlation Coefficients

Specification	Model	γ_0	τ^2_0	π_0	$\text{var}(P_j)$	ICC
Model 1	Empty model (constant)	- 0.3731	0.8454	0.4078	0.0493	0.2044
Model 2	Empty model (constant)	- 0.5631	0.4947	0.3628	0.0264	0.1307
	Empty model (region)	- 0.5631	0.2364	0.3628	0.0126	0.0671

5.1 Empty-model results. Intraclass correlation coefficients

In Table 4 and 5, parameters from empty model analysis to calculate the intraclass correlation coefficient (ICC) are presented. γ_0 represents the constant of the regression, τ^2_0 is the random effect of the regression, π_0 is the logistic transformation of the constant and $\text{var}(P_j)$ denotes the variance of the probability of producing a cesarean section in the hospital j . In model 1, intraclass correlation coefficient accounts for 20.44% of the variance in the sample. In other words, one fifth of variability of the probability of developing a cesarean section is due to differences between hospitals.

In model 2 a variance decomposition specification shows the inter-hospital variance attributable to the region effect. This model finds that almost 7% of the variance on the probability of undergoing a cesarean section is caused by the effect of region (approximately one third out of the hospital variance), whereas the 13.07% of such variance is due to the region-independent hospital clustering.

5.2 *Random-intercept results*

In Table 5, random-intercept multilevel logistic regression results are displayed. *Supply induced demand* variable and hospitals located in the Caribbean Region are significant and positive in predicting the performance of cesarean sections. Public hospitals are less likely to perform cesarean sections than their private counterparts. Higher-income mothers and those with previous deliveries are more likely to undergo a cesarean section. Moreover, mothers who completed any kind of superior education show a strong association with undergoing a cesarean section. Accomplishment of cesarean section criteria yielded a positive effect on the probability of performing a cesarean section.

Table 5. Marginal Effects of Logistic Multilevel Model

Variables	Model 1				Model 2						
	Empty model		Random intercept model		Empty model		Random intercept model				
	n=11594		n=11594		n=11594		n=11594				
	Coefficient	(Standard error)	Coefficient	(Standard error)	Coefficient	(Standard error)	Coefficient	(Standard error)			
Hospital variables											
High complexity			0.0685	(0.2649)			0.0684	(0.2649)			
Supply induced demand			0.0459	** (0.023)			0.0459	** (0.023)			
Hospital located at Bogota region			1				1				
Hospital located at Central region			0.2263	(0.2385)			0.2262	(0.2386)			
Hospital located at Pacific region			0.5726	(0.4022)			0.5725	(0.4022)			
Hospital located at Eastern region			0.3610	(0.3444)			0.3609	(0.3444)			
Hospital located at Caribbean region			2.3513	*** (0.3490)			2.3511	*** (0.3490)			
Public hospital			- 0.5447	** (0.2585)			- 0.5447	** (0.2585)			
Teaching hospital			0.2463	(0.2242)			0.2463	(0.2243)			
Individual variables											
Mother less than 18 years old			- 0.1888	*** 0.0732			- 0.1888	*** 0.0732			
Mother between 19 and 35 years old			1				1				
Mother more than 36 years old			0.5984	*** (0.0694)			0.5984	*** (0.0694)			
Mother income			0.0714	*** (0.0137)			0.0714	*** (0.0137)			
Previous deliveries			0.3800	*** (0.0440)			0.3800	*** (0.0440)			
Users per contract			- 0.0731	(0.0666)			- 0.0731	(0.0666)			
Female			0.0255	(0.0411)			0.0255	(0.0411)			
Unknown mother education			0.0775	(0.0506)			0.0775	(0.0506)			
Mother education less than completed primary			1				1				
Mother completed primary education			- 0.0043	(0.0845)			- 0.0043	(0.0845)			
Mother completed secondary education			0.0463	(0.0591)			0.0463	(0.0591)			
Mother completed any kind of superior education			0.3053	*** (0.1073)			0.3053	*** (0.1073)			
Dependent worker			1				1				
Self-employed worker			0.1383	* (0.0717)			0.1383	* (0.0717)			
No working			- 0.2227	(0.2159)			- 0.2227	(0.2159)			
C-section criteria accomplished			0.9284	*** (0.1147)			0.9284	*** (0.1147)			
Constant	- 0.3731	*** (0.1433)	- 1.0103	*** (0.3286)	- 0.5631	*** (0.1316)	- 1.0102	*** (0.3287)			
Variance											
Random effect of constant	0.8454	(0.1180)	0.4528	(0.0789)	0.4947	(0.1391)	0.4529	(0.0790)			
Random effect of region					0.2364	(0.0572)	0.0851	(0.0310)			
				Wald Chi2 (Prob > Chi2)= 367.43 (0.0000)				Wald Chi2 (Prob > Chi2)= 367.42 (0.0000)			
* Significance at 10.00%.											
** Significance at 5.00%.											
*** Significance at 1.00%.											

6 Discussion

Applying multilevel analysis to mothers belonging to a single insurer in Colombia, clustering from both hospitals and region has been demonstrated. Roughly 20% of variance is attributable to the gathering effect of medical facilities. Nonetheless, through an alternative specification of the error term, we find that one third of such variance is attributable to the effect of the geographic area. Focusing on the causes of these clustering effects according to Phelps and Money's approach (1993), some hypothesis are suggested. Differences in the independent variables between both specifications have not been found confirming consistency between both models.

On the demand side, it is possible that mothers' preferences play a key role in the decision of undergoing a cesarean section. An extensive amount of literature has been published demonstrating the effect of mother preferences on the performance of cesarean section (Blomqvist and Christiansen, 2004; Hsu *et al.*, 2008; Toll anes *et al.*, 2008). In this analysis, *Region* might be capturing a systematic effect on mothers' preferences, for instance, a cultural-related acceptance or rejection to cesarean section. Regarding region, this analysis reveals that mothers attended in the Caribbean Region are more likely to undergo cesarean sections. Despite that the fact that we cannot isolate the supply effect, mothers' preferences might be shaping this effect; for this reason, further analyses isolating supply and demand effects are highly recommended. However, mothers' preferences become evident when observing that wealthier and more educated mothers are more likely to undergo a cesarean section. In the Colombian health system, there are no differential out-of-pocket payments on the choice between both procedures, such as copayments or fees. Therefore, on a background of information asymmetry between physicians and mothers with physicians having privileged information, mothers might make non-maximizing choices. On the one hand, higher-income mothers might perceive higher expected costs in terms of sick leaves, pain or merely "inconvenience" from undergoing a natural delivery, leading them to persuade the physician to perform a cesarean section (Barbosa *et al.*, 2003; Fabri and Murta, 2002; Tschudin *et al.*, 2009; Walker *et al.*, 2004). On the other hand, higher-income mothers might be more likely to undergo a cesarean section due to the "social inheritance effect" (Toll anes *et al.*, 2008). Because of higher-income mothers have a larger probability of being born by cesarean section during the former National Health System (because wealthier mothers often were attended under fee-for-service arrangements where the supply induced demand is characteristic), they might be more likely to prefer a cesarean section as well. Other aspects of demand like clustering of complicated deliveries among small regions must also be considered. Low socioeconomic conditions of a certain population covered by a given hospital might lead

to more complicated births and hence, to be more likely to end up in cesarean sections (Chhabra *et al.*, 2006).

On the supply side, hospital and physician factors might also be affecting the variation on the choice of the procedure of delivery. Hospitals often have clinical guidelines on the criteria of performing procedures. This might partially explain clustering of procedures at the hospital level demonstrating the relevance of considering the hospital as the unit of analysis. Concerning the region effects we also obtained, it is possible that hospitals belonging to the same area share clinical guidelines or experience; hence, they might cluster the likeliness of developing a procedure at the regional level. Incentives and hospital restriction policies might also be shared. Furthermore, in concordance with the clustering of complications on the demand side, hospitals within the same area and perceiving a given prevalence of complications, might bargain similar contractual arrangements with the insurer. Conversely, public hospitals appear to be more likely to perform vaginal deliveries. This trend might be consequence of reduced incentives for public hospital physicians to perform cesarean sections.

Physicians' factors might also generate clustering on the choice of procedures. For example, it is possible that physicians do not experience a substantial degree of mobility during their lives; allowing them to share experiences with colleagues practicing in the same region and leading them to exhibit "school effects" on the performing of procedures (Epstein and Nicholson, 2009). Future research on physician mobility and school effects in Colombia is recommended to answer these questions. Another plausible factor is the risk of being sued. For example, physicians taking care of higher-income or more educated mothers could perceive a higher risk of facing legal actions if any complication arises during delivery leading them to carry out "defensive choices" (Vimercati *et al.*, 2000). *Supply induced demand* is a variable determining positively the likeliness of performing cesarean sections. Even though, this variable might generate cluster effects at the hospital level, it is also feasible that, as hospitals share contractual arrangements with insurers, they may also share the physician wage structure, leading to clustering effects at the region level.

Because we assess data from a single insurer, these results should be carefully inferred when considering other populations. We strongly suggest the inclusion of useful variables for small-area variation analysis in future national representative health surveys.

7 Conclusion

In this paper, a multilevel analysis (using data from a private Contributory Regimen insurer) is used to estimate the clustering effects of hospitals on the performance of cesarean sections. Two models with different specifications are provided. The first estimates the whole variance attributable to the hospital. The second model presents a variance decomposition specification in order to estimate what fraction of the hospital-attributable variance is due to the geographic region the hospital is situated in.

In this research, regional effects on the probability of performing cesarean section are found. In addition, this paper goes further and finds clustering effects at the health facility level, which is a more intuitive unit of analysis than regions. Some important findings are obtained. First, both the region and the health facility explain a considerable fraction of the variance on in performing cesarean sections. Unlike the previous literature, this paper proves that the effect of hospitals are relevant and that only a fraction of this effect can be explained by the region, challenging preceding approaches in which the geographical area comprises the unit of analysis. Moreover, this paper uses hierarchical methods, which may be more appropriate since they fit the data structure better than earlier linear techniques. We also find that the mother's income and any completion of superior education are significantly more likely to correlate with having cesarean sections. Given that out of pocket payments do not play a role on the choice of the procedure, such a result reveals how patient preferences might actually affect the likeliness of undergoing a cesarean section.

These results show that taking steps to improve informational asymmetries and to increase the standardization of procedures is urgently needed for this insurer. Further research at the national level and with other medical procedures is highly recommended to figure out the extent of this problem in Colombia.

Regarding future research on small-area variation in health care, clustering by healthcare providers or insurers rather than broad geographical areas is suggested since sometimes it lacks an intuitive basis and ignores the effect of smaller units as the source of variation. We recommend the use of hierarchical rather than linear methods to estimate clustering effects because these techniques fit better the data and provide useful additional information.

8 Acknowledgements

Thanks to Adriana Camacho from Universidad de los Andes and Gilberto Barón from Universidad Javeriana for their invaluable comments during the presentation of this paper. Also I want to thank Carolina Lopera, Jesus Otero and Manuel Ramírez from Universidad del Rosario and Jorge Higinio Maldonado from Universidad de los Andes for their remarkable observations. Special thanks to Sharaf Khan for all his help. Comments are welcome and all errors are ours.

9 References

1. Althabe *et al.* 2006. Caesarean section rates and maternal and neonatal mortality in low-, medium- and high-income countries: an ecological study. *Birth* **33** (4):270-277.
2. Bardey D, Lesur, R. (2006) Optimal regulation of health system with induced demand and “ex-post” moral hazard. *Annales d'Economie et de Statistique* **83-84** : 279-293.
3. Bech M, Lauridsen J. 2007. Dynamic Patterns and Small-Area Variation in Hospital Admissions. *iHEA 2007 6th World Congress: Explorations in Health Economics Paper*. Available at SSRN: <http://ssrn.com/abstract=992285>
4. Belizán JM, Althabe F, Barros FC, Alexander S. 1999. Rates and implications of caesarean sections in Latin America: ecological study. *British Medical Journal* **319**(7222): 1397-1402.
5. Blomqvist A. 1991. The doctor as double agent: information asymmetry, health insurance, and medical care. *Journal of Health Economics* **10**(4): 411-432.
6. Blomqvist S, Christiansen V. 2004. Taxation and heterogeneous preferences. *Working Paper # 1244*. Center for Economic Studies and Ifo Institute for Economics Research: Oslo, Norway. Available at URL: <http://SSRN.com/abstract=580543>
7. Castano R, Zambrano A. 2006. Biased selection within the social health insurance market in Colombia. *Health Policy* **79**(2-3): 313-324.

8. Chhabra P, Kumar AK, Tupil KA. 2006. Obstetric and neonatal outcomes in women who live in an urban resettlement area of Delhi, India: A cohort study. *Journal of Obstetrics and Gynaecology Research* **32**(6): 567–573.
9. Davis P, Gribben B, Scott A, Lay-Yee R. 2000. The “supply hypothesis” and medical practice variation in primary care: Testing economic and clinical models of inter-practitioner variation. *Social Science & Medicine* **50**(3): 407–418.
10. Epstein A, Nicholson, S. 2009. The formation and evolution of physician treatment styles: an application to cesarean sections. *Journal of Health Economics*. In press: doi:10.1016/j.jhealeco.2009.08.003
11. Fabri RH, Murta EFC. 2002. Socioeconomic factors and cesarean section rates. *International Journal of Gynecology & Obstetrics* **76**(1): 87-88
12. Flamm BL, Quilligan EJ. 1995. *Cesarean Section: Guidelines for Appropriate Utilization*. New York, NY: Springer-Verlag Inc.
13. Glover JA. 1938. The incidence of tonsillectomy in school children. *Proceedings of the Royal Society of Medicine* **31**: 95-112.
14. Grant D. 2009. Physician financial incentives and cesarean delivery: New conclusions from the healthcare cost and utilization project. *Journal of Health Economics* **28**(1): 244-250.
15. Gruber J, Kim J, Mayzlin D. 1999. Physician fees and procedure intensity: the case of caesarean delivery. *Journal of Health Economics* **18**(4): 473-490.
16. Grytten and Sørensen. 2003. Practice variation and physician-specific effects. *Journal of Health Economics*. **22**(3): 403-418.
17. Fuchs VR. 1978. The supply of surgeons and the demand for operations. *Journal of Human Resources*, **13** (Suppl: National Bureau of Economic Research conference on the Economics of Physician and Patient Behaviour, 1978): 35–56.

18. Hsu KH, Liao P, Hwang CJ. 2008. Factors affecting Taiwanese women's choice of cesarean section. *Social Science and Medicine* **66**(1): 201-209.
19. Iams JD, Chawla A. 1984. Patient costs in the prevention and treatment of post-cesarean section infection. *American Journal of Obstetrics and Gynaecology* **149**(4): 363-366.
20. Jaramillo H, Latorre C, Albán MC, Lopera C. 2008. El Hospital como organización de conocimiento y espacio de investigación y formación. Los recursos humanos en salud y su tránsito a comunidades científicas: el caso de la investigación clínica en Colombia. Centro Editorial Rosarista: Bogotá, Colombia.
21. Labelle R, Stoddart G, Rice T. 1994. a re-examination of the meaning and importance of supplier-induced demand. *Journal of Health Economics* **13** (3): 347-368.
22. MacDorman MF, Declercq E, Menacker F, Malloy MH. 2003. Infant and neonatal mortality for primary cesarean and vaginal births to women with “no indicated risk,” United States, 1998–2001 birth cohorts. *Birth: Issues in Perinatal Care* **33**(3): 175-182.
23. Mc Pherson K, Wennberg JE, Hoving OB, Clifford P. 1982. Small area variations in the use of common surgical procedures. *New England Journal of Medicine* **307**(21): 1310-1314.
24. Ministry of Social Protection (2000) “Resolución 412 de 2000”.
25. Ministry of Social Protection, Pan American Health Organization. 2007. *Situación de salud en Colombia: indicadores básicos*. Ministerio de la Protección Social: Bogotá, D.C.
26. Newhouse JP, Phelps CE. 1976. New Estimates of Price and Income Elasticities for Medical Care Services. In *The Impact of Health Insurance on the Health Services Sector*, Rosett R (ed). National Bureau of Economic Research: New York.
27. Pauly M, Satterthwaite M. 1981. The pricing of primary care physicians' services: a test of the role of consumer information. *Bell Journal of Economics* **12**(2): 488–506.
28. Phelps CE, Mooney C. 1993. Variation in the use of medical practice use: causes and consequences. In *Competitive Approaches to Health Care Reform*, Arnould RJ, Rich RF, White WD (eds). Urban Institute Press: Washington, D.C.

29. Phelps CE, Mooney C, Mushlin AI, Handy B, Perkins N. 1994. Doctors have styles - And they matter! *Working Paper # 2057*. University of Rochester Libraries: Rochester, NY. Available at URL: <http://hdl.handle.net/1802/2057>
30. Phelps CE, Newhouse JP. 1974. Coinsurance and the demand for medical care. *Review of economics and statistics* **56**(3): 334-342.
31. Phelps CE, Parente S. 1990. Priority setting and medical practice assessment. *Medical Care* **28**(8): 703-723.
32. Raudenbush SW. 1989. The analysis of longitudinal, multilevel data. *International Journal of Educational Research* **13**(X): 721-740.
33. Rabe-Hesketh S, Skrondal A. 2008. *Multilevel and longitudinal modeling using Stata*. Stata Press: College Station, TX.
34. Roos NP, Flowerdew G, Wajda A, Tate RB. 1986. Variations in physician hospital practices: a population-based study in Manitoba, Canada. *American Journal of Public Health* **76**(1): 45-51.
35. Ruiz A, Gómez C, Londoño D. 2001. *Investigación clínica: Epidemiología clínica aplicada*. Centro Editorial Javeriano: Bogotá D.C., Colombia
36. Snijders TAB, Bosker RJ. 1993. Standard errors and sample sizes for two-level research. *Journal of Educational Statistics* **18**(3): 237-259.
37. Snijders TAB, Bosker RJ. 1999. *An introduction to basic and advanced multilevel modeling*. The Cromwell Press: Trowbridge, UK.
38. Stano M. 1991. Further issues in small area variations analysis. *Journal of Health Politics, Policy and Law*. **16**(3):573-588.
39. Stano M. 1993. Evaluating the policy role of the small area variations and physician practice style hypotheses. *Health Policy* **24** (1): 9-17.

40. Tita ATN *et al.* 2009. Timing of elective repeat cesarean delivery at term and neonatal outcomes. *New England Journal of Medicine* **360**(2): 111-120.
41. Tollånes MC, Rasmussen S, Irgens LM. 2008. Cesarean among relatives. *International Journal of Epidemiology* **37**(6): 1341-1348.
42. Tschudin S, *et al.* 2009. Pregnant women's perception of cesarean section on demand. *Journal of Perinatal Medicine* **37**(3): 251–256.
43. Vecino-Ortiz AI. 2008. Determinants of demand for antenatal care in Colombia. *Health Policy* **86**(2-3): 363-372.
44. Villar J *et al.* 2006. Caesarean delivery rates and pregnancy outcomes: the 2005 WHO global survey on maternal and perinatal health in Latin America. *Lancet* **367** (9525):1819-1829.
45. Vimercati A *et al.* 2000. Choice of cesarean section and perception of legal pressure. *Journal of Perinatal Medicine* **28**(2): 111-117.
46. Wennberg JE. 1985. On patient need, equity, supplier induced demand, and the need to assess the outcome of common medical practices. *Medical Care* **23** (5): 512–520.
47. Wennberg J, Gittelsohn A. 1973. Small area variations in health care delivery. *Science* **182**(117): 1102–1108.
48. World Health Organization. 1985. Appropriate technology for birth. *Lancet* **2**(8452): 436–437.