MEASURING OVERTREATMENT: A STRUCTURAL MODEL TO ESTIMATE THE IMPACT OF NON-CLINICAL FACTORS ON HEALTHCARE UTILIZATION

by

ALEJANDRO ARRIETA

A dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Economics

written under the direction of

Professor Roger W. Klein

and approved by

New Brunswick, New Jersey

October, 2008

ABSTRACT OF THE DISSERTATION

MEASURING OVERTREATMENT: A STRUCTURAL MODEL TO ESTIMATE THE IMPACT OF NON-CLINICAL FACTORS ON HEALTHCARE UTILIZATION

by ALEJANDRO ARRIETA

Dissertation Director:
Professor Roger W. Klein

It is well acknowledged that, in the agency relationship between physicians and patients, the informational advantage gives doctors an incentive to deviate from the appropriate treatment, thus incurring over- or under- utilization. However, the empirical consequence of this problem has not been adequately considered. In particular, physician agency creates a gap between appropriate treatment and actual treatment whose characteristics and effects on estimation are analogous to a classification error.

This thesis proposes a structural model based on misclassification in which the physician behavior characterizes the structure of the measurement error. The model produces consistent estimators and is able to measure the degree of over- and under-utilization by separating out the effect of clinical and non-clinical variables on treatment decision. The model is applied to cesarean section deliveries performed in New Jersey in 1999-2002. The results show a moderate but growing rate of non-clinically required c-sections of around 3.2%, implying that the rapid growth of c-section rates over these years is explained mainly by non-clinical factors.

In the second chapter, the model is used to study how reform in the Peruvian health system has increased physician incentives to overuse c-sections in private hospitals. C-section rates in the private sector grew from 27% to 48% after the health reform of 1997, while the rates remained constant at 19% in the public sector. Using a national survey, it is estimated that each year more than 13 thousand women are over-treated, having a c-section without medical reasons. This document highlights the consequences of unnecessary c-sections on women's reproductive rights, and establishes important implications and recommendations for other health reforms in Latin America.

The third chapter extends the parametric estimation of the structural misclassification model to a semi-parametric estimation based on a double-index semi-parametric maximum likelihood with bias correction (Klein and Vella, 2008). I show that misspecification error due to a wrong assumption in error distribution may lead to an important inconsistency in parametric estimates, thus justifying the use of a semi-parametric technique to support results. The parametric and semi-parametric models are compared using a Monte Carlo study and an application for c-section deliveries.

ACKNOWLEDGEMENT

This thesis is the result of a long-term effort that was constantly guided for my advisors Roger Klein and Louise Russell. To them I am grateful because after discussions and revisions, the idea of a paper has now become a concrete dissertation. Only with the perspective of years working together, I can recognize now that their guides were also inspiration and motivation to continue with this project. From them I learned not only to succeed this thesis, but also to strength my professional development.

Mark Killingsworth, Derek DeLia, and Alan Monheit were also highly valuable with suggestions and revisions for this dissertation. I am also thankful to participants in the microeconomic seminar at Rutgers University, the health seminar at UMDNJ, and the Health Economics interest group workshop at AcademyHealth.

Finally, my acknowledgement also goes to the Center for State Health Policy at Rutgers University for data availability, to the Institute for Health, Health Care Policy and Aging Research at Rutgers for financial support through the Excellent Research Fellow, to the Economic and Social Research Consortium for the grant ACDI-IDRC 2006-PM 55 to study the cesarean section problem in Peru, and to the Inter-American Development Bank for financial support through a Research Fellow in the last stage of my research.

DEDICATION

To my wife Karin, my son Alonso and my daughter Natalia for their infinite support and love, and for their capacity to give up hours in family for allowing me to complete a personal goal. To my parents to believe in the returns to education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Over-utilization of medical procedures is common in US medicine, and it is considered a threat to health care quality and cost. Defined as the provision of health care service under circumstances in which its potential for harm exceeds the possible benefit (Chassin, et al., 1998), overuse has been widely identified in drug prescriptions and surgeries. The largest study of overuse was performed under the RAND health services utilization study. Based on a representative sample of Medicare beneficiaries, RAND researchers found that on average only 63% of cases in a set of procedures (coronary angiography, gastrointestinal endoscopy, carotid endarterectomy, coronary artery bypass graft, percutaneous transluminal angioplasty, and hysterectomy) were rated appropriate (Sharpe and Faden, 1996).

Methodologically, the definition of over-utilization relies on how appropriate treatment is defined. In some cases, appropriateness is rated on a case-by-case level. Some examples include the RAND-modified Delphi approach where a panel of physician experts rate individual cases for a specific procedure. In other cases, appropriateness is based on averages, and it reflects public health policy recommendations based on countries' experiences. Some examples are recommended percentages for cesarean deliveries or mortality rates for specific treatments guided by Healthy People or WHO.

There are some problems with both approaches. On one hand, the greater flexibility of a case-by-case approach is costly and open to criticisms related to subjectivity and inaccuracy. On the other hand, recommended percentages or rates are too rigid and unable to adapt to changes in demographics and new advances in medicine. An alternative approach is to use risk-adjustment models to define appropriateness based on a set of health status variables. Risk-adjustment has been broadly used to measure health care quality and performance, and hospital payment (Iezzoni, 2003). One important problem with risk-adjustment models is the use of observed ex-post health outcomes to define appropriateness. If, as is known, health outcomes are affected not only by clinical but also by non-clinical factors, then outcomes should also be adjusted for these non-clinical variables. There are many ways in which non-clinical factors may affect health outcomes. In particular, the health economics literature shows how the informational advantage of physicians regarding patients' health status creates incentives to overuse or underuse medical procedures according to specific physician's objectives.

Ever since Arrow's paper on uncertainty in the healthcare market (Arrow, 1963), the informational inequality in the doctor-patient relationship has become an issue that is now well identified (McGuire, 2000). An important consequence of this informational asymmetry is the physician-induced demand theory (Fuchs, 1978; Dranove, 1988), where doctors may exert influence over patients and intentionally shift the patient demand curve, thus increasing health care services against patients' best interest.

Also related to this problem, the health service research has placed a growing attention on explaining the large variation of utilization rates across geographic areas (Wennberg, 2002; Fuchs et al., 2001). There is strong evidence that these regional disparities are related to non-clinical factors such as fear of litigation, racial biases, socioeconomic differences, and institutional differences among others.

In all these cases is important to find the clinical and non-clinical factors that define an appropriate treatment. However, some studies recognize that observed treatments may be “contaminated” with the effect of non-clinical factors. As the patient-physician relationship involves a particular interaction between the patient’s health status and the physician’s incentives, controlling for non-clinical factors by itself is not enough to correct the problem.

The main contribution of this document is a methodology based on a structural misclassification model to estimate the impact of non-clinical factors on healthcare utilization. While the doctor sees which delivery method is warranted by the patient’s health status, this “true” response is not seen by the econometrician. If financial incentives are strong enough to overcome professional ethics, the doctor will influence a patient to have a treatment even though it is not clinically necessary. In this case, the appropriate choice is affected by the physician’s decision, resulting in a “misclassified” outcome. Since this problem is analogous to the misclassification problem, it also shares its consequences. In general, measurement error on limited dependent variables (misclassification) leads to biased and inconsistent estimators. However, in this particular structure, the classification error will not be random, but a behavioral model can be used to incorporate the structure of the measurement error into the estimation process in order to find consistent estimators.

In the first chapter I develop a methodology based on a structural misclassification model to estimate the impact of non-clinical factors on healthcare utilization. That chapter describes the structural misclassification model and establishes a parametric solution. A Monte Carlo study is then used to compare the effect on estimators' consistency of four different approaches to estimate risk-adjusted utilization rates, ranging from not considering the misclassification problem to considering it adequately. Finally, this chapter provides an application to the case of cesarean section deliveries in New Jersey.

In the second chapter I apply the structural misclassification model to study over-treatment of c-sections in Peru. The reform of the Peruvian private health system has increased doctor's financial incentives to overuse medical procedures. After the reform, managed healthcare organizations got enough market power to push healthcare prices down, but were unable to increase the access to private health. With lower incomes and a stagnant market, doctors have higher incentives to increase cesarean sections because they are more profitable than vaginal deliveries. This explains why, after the reform, the c-section rate in the private sector has grown from 27% to 48%, but this rate reaches 66% when women have access to private insurance. In the public sector, however, the c-section rates remain in around 19%. This study is relevant because it identifies the determinants of over-use of c-sections and estimates their effects. Based on the results this chapter defines policy recommendations oriented to reduce the problems in a more effective way.

Finally, the third chapter explores the misspecification error in a parametric structural misclassification model of over-treatment. In this special case, the model reduces to a bivariate probit with partial observability. Through a Monte Carlo study I found that misspecification creates an important bias that cannot be reduced even with larger sample size. Motivated by this result, a semi-parametric estimation is suggested and evaluated. Because of its special characteristics, this model of over-treatment can be estimated by double-index semi-parametric maximum likelihood estimation with partial observability. The results of both models, the parametric and the semi-parametric model, are compared using a Monte Carlo study and an application with data on cesarean sections in New Jersey.

CHAPTER 1: A Structural Misclassification Model to Estimate the Impact of Non-Clinical Factors on Healthcare Utilization

1.1 Introduction

Although the fact that physician incentives affect health care utilization rates is known, it has not been adequately considered in empirical literature. The usual approach has been to estimate binary dependent models (usually logit or probit) and control for clinical and non-clinical factors depending on the variables of interest. However, this approach does not consider that physician incentives break the correspondence between appropriate treatment and observed treatment. The definition of appropriate treatment is based on patient health status only, the latter being only observed by the physician. When incentives are strong enough, the physician deviates from appropriate treatment and, therefore, the observed treatment will not reflect the clinical characteristics of the patient. In that regard, this problem can be seen as a misclassification problem where the measurement error in the binary dependent variable is proportional to the strength of physician incentives to deviate from appropriate treatment.

Two examples are illustrative. The first is related to racial differences in healthcare access. One of the most studied cases is lower access to cardiovascular procedures in African Americans (Kressin and Peterson, 2001; Ford and Cooper, 1995; Van Ryn and Burke, 2000). In this case, an African American patient with a poor health condition requires a cardiovascular surgery. Based on health status, the appropriate treatment -observed only by the doctor- should be the utilization of the procedure. However, if the doctor has a racial bias, he may influence the treatment choice by omitting to suggest the surgery. In that case, the observed outcome will be "misclassified" resulting in under-utilized cardiovascular procedures for African Americans.

However, some authors suggest that, in the case of cardiovascular surgeries, the problem is not under-use for African Americans, but over-use among white patients (Schneider et al., 2001). This means that based on additional non-clinical factors, the doctor may also influence treatment choice for white patients by suggesting the surgery when it is not required. In this case, both over- and under-healthcare utilization coexist, and the observed outcome will thus be "misclassified" in both directions: observation of surgery when it was not required, and no observation of surgery when it was required.

The second example is related to the demand inducement theory, where fee-for-service pricing creates financial incentives for the physician to recommend unnecessary medical procedures. A well known case is cesarean section delivery (Gruber and Owings, 1996; Das, 2002; Tussing and Wojtowycz, 1993). While the doctor sees which delivery method is warranted by the woman's health status, this "true" response is not seen by the econometrician. If financial incentives are strong enough to overcome professional ethics, the doctor will influence a woman to have a c-section even though it is not clinically necessary. In this case, the appropriate choice is affected by the physician's decision, resulting in a "misclassified" outcome that is identified by the econometrician. Note that in this example, "misclassification" runs in one direction only: observation of c-section when it is not required. This happens because there is no monetary incentive to perform a vaginal delivery, and, to top it off, there is fear of litigation. These factors result in a strong disincentive to avoid a vaginal delivery when a c-section is required.

Since this problem is analogous to the misclassification problem, it also shares its consequences. In general, measurement error on limited dependent variables (misclassification) leads to biased and inconsistent estimators. When misclassification is not adequately corrected, it understates parameter estimates, and overstates standard errors. Literature has focused mainly on the case where misclassification is originated randomly by errors in the report or record of a categorical variable (Hausman et al., 1998; Magder and Hughes, 1997; see Kenkel et al., 2004 for an application). Additionally, Abrevaya and Hausman (1999) and Lewbel (2000) have considered the case in which misclassification depends on some covariates, imposing strong conditions for identification. However, there are cases where a decision maker is able to alter the true outcome. When this happens, the classification error will not be random, but a behavioral model can be used to incorporate the structure of the measurement error into the estimation process in order to find consistent estimators.

In this chapter I develop a methodology based on a structural misclassification model to estimate the impact of non-clinical factors on healthcare utilization. This brings three improvements compared to previous literature: First, I obtain consistent estimates for clinical and non-clinical characteristics. Second, I model physician's behavior and its interaction with patient's health status. Third, I am able to estimate the rate of inappropriate treatments defined as those related to non-clinical factors (misclassification probability) and the risk-adjusted utilization rate based only on health characteristics (removing non-clinical factors).

The second section of this chapter describes the structural misclassification model and establishes a parametric solution. In the third section, I use a Monte Carlo study to compare the effect on estimators' consistency of four different approaches to estimate risk-adjusted utilization rates, ranging from not considering the misclassification problem to considering it adequately. The fourth section provides an application to the case of cesarean section deliveries in New Jersey. The last section provides conclusions to the chapter.

1.2 A parametric estimation of structural misclassification

In the examples presented in the previous section, the physician observes the true health condition of his patient. Conditioned on the patient's health status, the doctor may deviate from appropriate treatment. This will happen if his personal goals overcome his professional ethics within his utility function. Therefore, the structure of this model is a simplified version of game theory models of inducement (De Jaeguer and Jegers, 2001; Xie et al. 2006). It is a simple version because I do not explicitly consider the patient's decisions. The physician's decision tree is shown in figure 1.1. In the first stage, nature determines a patient's state (healthy $h<0$, or sickly $h\geq 0$), and this can only be observed by the physician. Two treatments are considered: A and B. I assume that doctor's (monetary or non-monetary) benefits under treatment A are higher than treatment B. However, only one treatment is appropriate for each patient's health state.

In the second stage, the physician must choose the treatment based on his incentives (i). Based on the patient's health state, he may decide to perform the treatment that is appropriate for the patient ($i<0$) or an inappropriate one ($i\geq 0$) based on patient's health state (h). However, the physician will choose the treatment that gives him more utility in terms of monetary and non-monetary factors, and after discounting the intrinsic cost of acting against professional ethics. Therefore, the physician will choose treatment A or B that may or may not be the appropriate treatment for each patient's health state. For instance, a doctor may inappropriately recommend treatment A (cardiovascular surgery) for a healthy white patient and treatment B (no surgery) for a sickly African American patient. On the other hand, an obstetrician may inappropriately recommend treatment A (c-section) for both, risk and non-at-risk pregnant women.

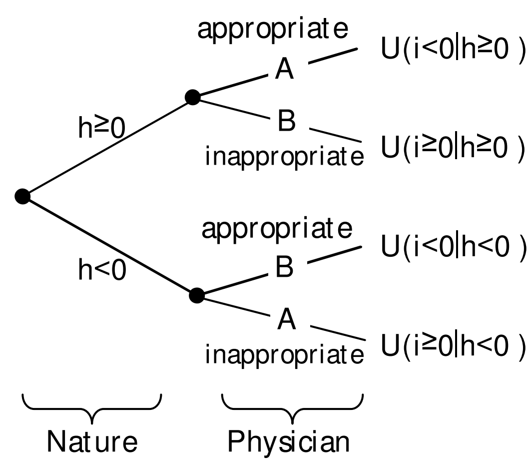

Figure 1.1: Physician's decision tree

Notice that a more complete model may explicitly include a third stage where the patient accepts or rejects the treatment based on obtained medical information (Xie et al. 2006). However, figure 1.1 may be seen as a one-step backward induction in which physicians take patients' actions into consideration. In that regard, index i in the utility function represents incentives that are net of the cost implied by acting against the doctor's ethical standards, but also the expected loss that would occur when a patient that is well informed decides to leave the doctor. The expected loss depends on how well the physician knows his patient (that is, if he knows the patient could collect medical information on his own, as well as their degree of resistance to physician's influence).

The econometric model is described following figure 1.1. In the first stage, a patient's health status (h) is determined by a set of observable clinical characteristics or risk factors (x) and unobserved risk factors (ε_h) .The physician can observe a patient's health condition:

$$h = x\beta + \varepsilon_h \tag{1.1}$$

This is the health status equation. There are two possible treatments, $\tilde{y} = \{0,1\}$. The patient will require treatment $\tilde{y} = 1$ if health status exceeds zero

$$\tilde{y} = 1 \text{ if } h = x\beta + \varepsilon_h \geq 0$$
$$\tilde{y} = 0 \text{ otherwise}$$

The econometrician observes the doctor's treatment choice y but not $\tilde{y}$. Without physician incentives to alter the required treatment, $y = \tilde{y}$ and in that case any binary model estimation (logit or probit) will be consistent because the probability of observing the treatment choice and the probability of not observing it are respectively

$$\Pr(y = 1) = \Pr(\tilde{y} = 1) = \Pr(h \geq 0)$$
$$\Pr(y = 0) = \Pr(\tilde{y} = 0) = \Pr(h < 0)$$

However, if the physician decides to do the surgery when it is not needed -like in the cesarean section case- or not to do the procedure when it is required –like in the case of cardiovascular surgery- then the binary model estimation will be inconsistent because $\Pr(y=1) \neq \Pr(\tilde{y}=1)$. In those cases the econometrician observes a "misclassified" treatment.

In the second stage, the physician decides to alter the required treatment based on his incentives (i). Incentives depend on doctor's characteristics and patient's characteristics that are observed (z) and unobserved (ε_i) by the doctor.

$$i = z\gamma + \varepsilon_i \tag{1.2}$$

This is the physician incentives equation. When incentives equal or exceed a threshold 0, the doctor may proceed, depending on health status, with the inappropriate treatment, thus altering the appropriate choice with probabilities (misclassification probabilities)

$$\alpha_0 \equiv \Pr(y=1 \mid \tilde{y}=0) = \Pr(i \geq 0 \mid h < 0) \tag{1.3a}$$

$$\alpha_1 \equiv \Pr(y=0 \mid \tilde{y}=1) = \Pr(i \geq 0 \mid h \geq 0) \tag{1.3b}$$

α_0 is the probability of over-utilization (doctor performs a surgery when it is not required). α_1 is the probability of under-utilization (doctor does not perform a surgery when it is required). These two probabilities define the probability of observing the surgery:

$$\begin{aligned} \Pr(y=1) &= \Pr(i < 0 \mid h \geq 0)\Pr(h \geq 0) + \Pr(i \geq 0 \mid h < 0)\Pr(h < 0) \\ &= \alpha_0 + (1 - \alpha_0 - \alpha_1)\Pr(h \geq 0) \end{aligned} \tag{1.4}$$

Note that the second equality in equation 1.4 corresponds to equation 5 in Hausman et al. (1998). Clearly, without physician incentives, $\alpha_0 = \alpha_1 = 0$ and therefore $\Pr(y=1)$ collapses to $\Pr(h \geq 0) = \Pr(\tilde{y}=1)$, returning the consistent estimation of binary models. For the cardiovascular example, it is expected that $\alpha_0 > 0$ when the patient is white and $\alpha_1 > 0$ when the patient is an African American. This is the case when misclassification runs in both directions. In this case the probability of observing a patient with ($y=1$) or without ($y=0$) cardiovascular surgery is respectively:

$$\Pr(y=1) = \Pr(i < 0 \mid h \geq 0)\Pr(h \geq 0) + \Pr(i \geq 0 \mid h < 0)\Pr(h < 0)$$
$$\Pr(y=0) = 1 - \Pr(i < 0 \mid h \geq 0)\Pr(h \geq 0) - \Pr(i \geq 0 \mid h < 0)\Pr(h < 0)$$

For the cesarean section example, I expect over-utilization ($\alpha_0 > 0$) because of the disincentives to proceed with a vaginal delivery when a c-section is needed. It means that misclassification runs in only one direction given that $\alpha_1 = 0$. In this case the probability of observing a c-section ($y=1$) and the probability of a vaginal delivery ($y=0$) are respectively:

$$\Pr(y=1) = \Pr(h \geq 0) + \Pr(i \geq 0 \mid h < 0)\Pr(h < 0)$$
$$\Pr(y=0) = \Pr(i \leq 0 \mid h < 0)\Pr(h < 0)$$

Haussman et al. (1998) and Lewbel (2000) discuss the conditions for identification of misclassification models. In particular, the *monotonicity condition* (MC) is required to identify the parameters. For the general misclassification model, the MC is $\alpha_0 + \alpha_1 < 0$. In terms of the problem, this condition is generally satisfied because the degree of physician incentives is relatively small given the punishment in terms of reputation and lawsuits that result if inappropriate treatments are very high. Note that in the case of misclassification in one direction (either over- or under- utilization), the MC is automatically satisfied since there is only one probability, which I safely assume is below 1.

The parameters β and γ in equations 1.1 and 1.2 can be estimated with MLE, maximizing the likelihood function

$$L(\beta,\gamma)=\prod_{i=1}^{n}\Pr(y=1)^{y}\Pr(y=0)^{1-y} \qquad (1.5)$$

Notice that if the errors $\varepsilon_h,\varepsilon_i$ are not independently distributed (with correlation ρ), the problem becomes a bivariate model. Additionally, if it is assumed that the error terms are normally distributed, the problem becomes a bivariate probit (Amemiya, 1985). For the general case described by equation 1.4, the likelihood function is

$$L(\beta,\gamma;\rho)=\prod_{i=1}^{n}[2\Phi_2(x\beta,-z\gamma;\rho)]^{y}[1-2\Phi_2(x\beta,-z\gamma;\rho)]^{1-y} \qquad (1.6)$$

Where Φ_2 is standard bivariate normal CDF. Notice that this model is a variety of Poirier's partial observability model (Poirier, 1980). However, there are important differences. The partial observability model considers two agents making decisions based on a common set of information. In this structural misclassification model, there is only one decision maker: the physician. The patient's health status is not a decision maker and consequently equations 1.1 and 1.2 are generally functions of two separate sets of variables, in contrast to the partial observability model. A variable may be in both equations if it carries information about the patient's health status and the physician's incentives (some examples are age, sex, weight, etc. depending on the analyzed treatment). Partial observability models have been used to address misspecification in simple probit or logit models (see for example Abowd and Farber, 1982), while in this document partial observability is obtained after adding structure to the misclassification problem. Finally, the likelihood function of the structural misclassification model is different to the partial observability model, but both models converges when misclassification run in only one direction (under- or over- utilization). Note that the loss of information due to limited observability reduces efficiency of the maximum likelihood estimator as in the partial observability models (Poirier, 1980; Meng and Schmidt, 1985).

The structural misclassification error model presented in this chapter rests on strong parametric assumptions. I have based the estimation on a bivariate probit, but the model can be easily extended to a bivariate logit. However, a natural extension is to get rid of the parametric assumptions and estimate this model semi-parametrically based on a multiple index model. This task is developed in Chapter 3.

1.3 Comparing methodologies: Monte Carlo simulation

In this section, I study the consequences of omitting or mis-specifying the impact of physician incentives on observed health care utilization. In particular, I study how consistency estimation is affected when I use the current methodology used in literature to estimate the influence of clinical factors on health outcome. The variable of interest is a dichotomous variable indicating whether the patient received the treatment under study or the alternative treatment (or no treatment). I consider the case in which physicians have incentives to alter the required treatment given the patient's health status. For simplicity and with the purpose to connect this study with the empirical application described in section 1.4, I focus on the case of physician incentives to over-utilize medical procedures.

In estimating the effects of clinical risk factors associated with specific treatments, the literature has followed two common approaches[1]: (i) Estimation of a simple binary (SB) model (logit or probit) with clinical factors as the only regressors, and (ii) Estimation of a simple binary model with controls (SBC), where non-clinical factors are added to the SB model as control variables. Because the agency problem in the physician-patient relationship creates important interactions between a patient's health status and non-medical factors, the omission of non-clinical characteristics creates a serious omitted variable bias. When non-clinical factors are added as controls as in the SBC model, the bias is reduced but interactions are not appropriately captured, and the misclassification bias described in previous sections is not adequately corrected.

In that regard, I also study the following two approaches based on the structural misclassification model developed in section 1.2: (iii) Structural misclassification model with independent errors (SMCI), and (iv) Structural misclassification model where errors are allowed to be dependent (SMC). From an empirical perspective, the restriction of independent errors may be strong. Because there are many health characteristics that are non-observable to the econometrician but observable to the physician, the latter could

[1] To get rid of the potential "misclassification" problem, other approaches have considered the reviewing of medical records in light of professional guidelines based on clinical trials or expert opinions. Good examples of these approaches are the RAND guidelines and the American College of Cardiology/American Heart Association guidelines. For a discussion of the own problems of these methodologies see Leape et al. (2003).

make decisions based on information conveyed in ε_h. Given that some variables related to physician incentives are also unobserved by the econometrician, ε_i and ε_h may be correlated. In that regard, the error correlation also measures physician incentives that are difficult to observe because either the clinical factors do not define clearly what an appropriate treatment is or the physician's incentives go beyond measurable characteristics. Moreover, it is expected that the error correlation is negative for over-utilized procedures and positive for under-utilized procedures. Consider the case of the racial bias example described in sections 1.1 and 1.2. If the doctor observes health characteristics that cannot be easily captured by diagnosis codes or medical guidelines, but that imply a need to perform surgery on an African American patient, then it will be easier for a biased doctor to deviate from the appropriate treatment, and regarding any observable characteristic, the physician's incentive growths implying a positive error correlation. A similar argument can be used to expect a negative correlation for over-utilized procedures. For that reason, it is important to evaluate the consequence of imposing the independence restriction.

In order to assess the impact of these four approaches on estimator bias and consistency, I examine the results of Monte Carlo simulation. The true model representing equations 1.1 and 1.2 is

$$h = -1.5 + 0.5x_1 - x_2 + 2x_3 + \varepsilon_h$$
$$i = -2.5 - 1.5z_1 + z_2 + 0.5z_3 + 2z_4 + \varepsilon_i$$

Covariates x and z include dummy variables and continuous variables that were drawn from uniform distributions and trimmed chi-squared distributions to avoid outliers.

The error disturbances $\varepsilon_h, \varepsilon_i$, are drawn jointly from a bivariate standard normal distribution with correlation $\rho = 0.25$. For the design of the Monte Carlo study I consider 1000 independent random draws of a sample size of 5000. Table 1.1 reports the sample mean and standard error of parameters estimated over the 1000 draws.

Table 1.1: Monte Carlo simulations

	True parameter values	**Simple Binary Model (1)**	**Simple Binary Model with Controls (2)**	**Structural Misclassification Model with independent errors (3)**	**Structural Misclassification Model (4)**
		Clinical variables			
β_0	-1.50	-0.752	-1.340	-1.508	-1.502
		(0.044)	(0.074)	(0.092)	(0.089)
β_1	0.50	0.311	0.349	0.495	0.502
		(0.129)	(0.138)	(0.178)	(0.180)
β_2	-1.00	-0.544	-0.626	-0.985	-1.005
		(0.045)	(0.047)	(0.085)	(0.087)
β_3	2.00	1.249	1.408	1.972	2.003
		(0.068)	(0.072)	(0.114)	(0.117)
		Non-clinical variables			
γ_0	-2.50	–	–	-2.548	-2.505
				(0.215)	(0.211)
γ_1	-1.50	–	-0.510	-1.488	-1.503
			(0.143)	(0.321)	(0.321)
γ_2	1.00	–	0.301	0.989	0.995
			(0.046)	(0.127)	(0.126)
γ_3	0.50	–	0.176	0.493	0.498
			(0.066)	(0.144)	(0.145)
γ_4	2.00	–	1.069	1.978	2.005
			(0.048)	(0.127)	(0.126)
ρ	0.25	–	–	–	0.239
					(0.193)
Prob. approp. treatment †	0.249	–	–	0.244	0.250
				(0.015)	(0.015)
Prob. physic incentives ‡	0.112	–	–	0.100	0.110
				(0.018)	(0.020)

n=5000, 1000 simulations. Standard deviations in parentheses.

† Calculated using the marginal probability $\Pr(h \geq 0)$

‡ Calculated using the marginal probability $\Pr(i \geq 0)$

The results of the Monte Carlo simulation are consistent with the misclassification problem described by Hausman et al. (1998). In particular, the simple probit (SB and SBC models) understates the coefficients. The probit model with only patient's health related variables (SB) produces estimates that are biased downward by 35-50% (column 1). When physician incentives are added as control variables (SBC model), bias is still substantial in the case of health risk factors. Coefficient estimates of control variables (non-clinical characteristics) have a more substantial downward bias of around 45-70% (column 2).

Bias size in the health status and the non-clinical coefficients depends, among others, on two parameters: the error correlation and the degree of physician incentives (misclassification). Different Monte Carlo designs (not shown) were used to see the impact of both parameters on estimator biasedness. First, the bias increases in both sets of estimators when the error correlation gets closer to 1. Given the previous discussion related to the sign of the error correlation, this result implies that bias will be larger in the case of under-utilized procedures than in the over-utilized procedures. It is also important to note that even in the case of small correlation, the bias does not vanish. Second, the bias in the health status estimators decreases and in the doctor's incentives estimators increases when the degree of incentives falls. A rough exercise shows that the bias in the health status estimators decreases almost proportionally with the reduction in the degree of incentives. On the other hand, a reduction in the degree of incentives increases the bias in the incentive equation more than proportionally when incentive is high, and less than proportionally when it is low.

Estimators of the structural misclassification models (SMCI and SMC) are consistent but present larger standard errors (columns 3 and 4). Compared with these models, the simple probit (SB and SBC models) overstates the precision of its estimates. As an implication for health care quality studies, the impact of health risk factors on utilization rates will appear to be less important than they really are when probit or logit is used. Confidence intervals will be narrower too.

There is a small discrepancy between the structural misclassification model estimators when error independence is erroneously imposed (column 3). In general, estimated coefficients of the SMCI model are biased by 1-2% in the case of risk factors and physician and patient characteristics (column 3). For the structural misclassification model that allows correlated errors (SMC model), the estimates are unbiased (less than 0.5%) and show slightly higher standard errors than the SMCI model (column 4).

An important feature of the structural misclassification model is that it allows us to neatly separate the estimated physician incentives probability and the utilization rate due to health status only. Table 1.1 reports the degree of doctor's incentives or percentage of misclassification. It was calculated as the marginal probability that physician incentives exceed threshold zero: $\Pr(i \geq 0) = \int \Pr(i \geq 0, h)dh$. The design of the Monte Carlo study implies a true degree of incentives of 0.112. The degree of incentives estimated by the SMCI model is downward biased by 11% (column 3), while the SMC model presents a smaller bias (column 4).

If the goal is to estimate risk-adjusted utilization rates, the appropriate measure that discards the effect of physician incentives on the "misclassification" bias will be the marginal probability of appropriate treatment. For this particular design where misclassification is related to over-utilization, the appropriate treatment is obtained only when health status exceeds zero, and therefore $\Pr(h \geq 0) = \int \Pr(i, h \geq 0) di$ will be the estimated utilization rate based on health status only, under the counterfactual that there are no physician incentives to influence the appropriate health outcome. For this Monte Carlo study, true probability of appropriate treatment is 24.9%. The SMCI model understates the probability by 1.7% (column 3). The bias is almost zero (less than 0.5%) when the probability of appropriate treatment is estimated using the SMC model (column 4).

1.4 An application to cesarean section deliveries in New Jersey

1.4.1 Data

This section applies the structural misclassification model to births in the state of New Jersey in the period 1999-2002. I use Hospital Patient Discharge Data collected by the New Jersey Department of Health and Senior Services. This data contains detailed information on each discharge from an acute care hospital including identification of the hospital, patient demographics and zip code of residence, diagnosis and surgical procedures classified by ninth revision of the International Statistical Classification of Diseases and Related Health Problems(ICD-9) codes and Diagnosis Related Group numbers (DRG), source of admission, and identification of payers. Additional socioeconomic information was collected from the US Census 2000, using the patient's zip code as the key variable for matching. Births were identified by DRG codes 370-375. Cesarean sections were identified by DRG 370-371 or ICD-9 code 74xx excluding 7491.

The selected sample includes women aged 15 to 49. I excluded deliveries performed in hospitals that in a particular year had less than 100 births (0.03% of the sample). Finally, I also exclude patients with wrong zip codes (0.5%) and patients with missing or wrong reported information in the variables of analysis (1.7%). The final sample used for the estimation considers a total of 403,660 women.

This estimation uses two sets of variables. A first set comprises women's health characteristics identified according to diagnosis codes. I follow previous health service research to select the clinical variables that were seen as more relevant to explain c-sections (Keeler et al., 1997; Aron et al., 1998; DiGiuseppe et al., 2001; and Rahnama et al., 2006). The second set of variables comprises patient and physician characteristics that may drive doctor's incentives. The complete list of variables and their mean values for vaginal and c-section deliveries are reported in Table 1.2.

Table 1.2: Sample mean of health and non-health related variables: New Jersey 1999-2002 (Percentages unless noted)

Variable	Vaginal Delivery	Cesarean Section	Full Sample
Cesarean delivery	0.00	100.00	24.83
Clinical Variables			
Age (years)	28.62	30.47	29.08
Previous cesarean delivery	4.01	42.00	13.44
Multiple gestation	0.71	2.84	1.24
Admission by emergency	5.58	3.30	5.02
Long labor	0.76	0.78	0.77
Elderly primigravida >=35 y.o.	0.69	1.39	0.86
Breech or transverse lie presentation	2.40	22.98	7.51
Diabetes	3.45	5.53	3.97
Hypertension	3.24	3.57	3.33
Pre-eclampsia	1.62	1.42	1.57
Oligohydramnios	0.21	0.21	0.21
Polyhydramnios	0.30	0.94	0.46
Abruptio placenta	0.47	0.60	0.50
Full or partial placenta previa	0.10	0.87	0.29
Patient and Physician related variables			
Woman is married	65.90	71.48	67.28
Woman is full time employed	34.85	40.46	36.24
White (non-Hispanic)	43.43	44.16	43.61
Black (non-Hispanic)	13.19	12.14	12.93
Hispanic	17.40	17.46	17.42
Zip code mean household income (thousands, $)	56.00	57.15	56.29
Patient payment (uninsured)	8.22	6.41	7.77
Medicaid payment	10.62	8.60	10.11
HMO payment	57.04	58.01	57.28
Yearly average of births in Hospital (thousands)	2.48	2.61	2.51
Obs&Gyn Physician	89.67	91.09	90.02
Number of observations			
Total	303,434	100,226	403,660
Year 1999	76,610	22,193	98,803
Year 2000	77,571	24,371	101,942
Year 2001	76,273	26,027	102,300
Year 2002	72,980	27,635	100,615

1.4.2 Structural Model of Physician Behavior

I use the structural misclassification model described in section 1.2 to measure the probability of unnecessary c-sections and to estimate the probability of appropriate cesarean section rates by removing non-clinical factors. However, it is important to highlight that the physician incentive equation may capture patient's choice rather than physician's influence exclusively. Even though c-sections by women's choice are not allowed in the USA and professional guidelines consider c-sections for non-medical reasons "*to fall outside the bounds of best professional practice*"[2], c-sections related to non-clinical factors are a possibility. The lack of variables related to patient's choice made impossible to identify these two non-clinical factors. Notice that this model can be easily extended to test for physician-induced demand (PID), however the data collected for this application does not allow us to identify PID because of lack of an exogenous shock on medical income on that period.

Again, in the case of c-sections the misclassification runs in only one direction: over-utilization. Assuming standard normal distributions for the error terms, the model is estimated by MLE using the likelihood function described in equation 1.6 that in this case becomes:

$$L(\beta,\gamma;\rho)=\prod_{i=1}^{n}[1-\Phi_2(-x\beta,-z\gamma;\rho)]^y[\Phi_2(-x\beta,-z\gamma;\rho)]^{1-y}$$

where the probability of observing a c-section is:

$$\Pr(y=1)=1-\Phi_2(-x\beta,-z\gamma;\rho)$$

---

[2] FIGO Statement on Cesarean Section. January 2007. http://www.figo.org/Cesarean.asp

In general, variables in the physician incentive equation may be classified in two types: First, variables observed by the doctor that signal the degree of patient-obtained medical information. These are mainly socioeconomic characteristics that can be observed directly (employment, marital status, etc.) or that can be inferred from patient population (patient ethnicity, patient geographic area residence, etc.) Usually, patients with lower socioeconomic status have less access to information, and also have less capability to use such information in the medical visit (Xie et al. 2006). To capture this effect I include four variables: (i) Social support, measured as the presence of a partner (married or life partner), may be perceived to improve the degree of information because decisions maybe taken on a couple basis. It is expected that it reduces physician incentives and therefore the probability of c-section. (ii) Woman's employment status may signal a potential compliance to c-sections. It is expected that fully employed women may prefer a c-section delivery because of the convenience in terms of scheduling and the lower pre-partum work. Consequently I expect higher physician incentives for full employed women. (iii) Ethnicity (a White, Black or Hispanic woman) may be perceived as a signal of access to medical information. It is expected that minorities have less access and worse use of medical information, which makes them more vulnerable to physician influence. (iv) Patient socioeconomic status is not easily observed by the doctor. Instead, the physician can infer the socioeconomic status from the zip code of the patient's residence. I include the average household income at zip code level. It is expected that a zip code with low income is perceived as lower socioeconomic status, thus increasing vulnerability to physician incentives (Pauly, 1980).

The second type of variables comprises factors that directly affect physician incentives. These are mainly institutional and contractual variables related to the physician itself or the health facility. I consider three variables: (i) The method of payment or patient's insurance condition (uninsured, Medicaid, HMO, other insurance). The type of insurance is important for doctor's incentives because it sets the method of payment. It is expected that the uninsured have the lowest rate of c-sections compared to non-HMO private insurance, since they must pay for the procedure. The capitation payment of an HMO reduces physician incentives, and so does prospective payment system under Medicaid. An additional factor that reduces incentives in the case of Medicaid patients is the low fees for obstetric procedures observed in New Jersey. (ii) The size of the hospital measured as the average yearly number of births observed in each hospital. It has been shown that hospital size has an impact of over-utilization due to the supply-sensitive service phenomenon (Wennberg, 2002). Larger hospitals usually have larger fixed costs increasing the incentive to use more expensive treatments to keep returns. Consequently, a higher probability of c-sections in larger hospitals is expected. (iii) The physician specialty (Ob/Gyn specialty) captures the tendency of over-utilization in more specialized doctors. It is expected to observe more c-sections when the delivery is attended by a specialist. Finally, year dummy variables are also included in the estimation of the physician incentive equation to capture the trend related to non-observed factors.

The set of variables related to clinical factors is not discussed here and I refer to the specialized health service literature for details (see section 1.4.1). However, it is important to highlight that all these variables are related to risk of pregnancy and delivery, and therefore they are expected to increase the probability of a c-section.

1.4.3 Results

Table A1.1 in the Appendix reports the estimation using three econometric methods: The simple binary model (SBM), the structural misclassification model with independent errors (SMCI), and the structural misclassification model (SMC). Compared to the SMC model, the simple probit (SBM) understates both the impact of woman's health characteristics and physician incentives (column 1). However, the bias is greater for incentive related variables, and that is explained by the small probability of non-clinically required c-sections (misclassification probability) observed in the data (estimated at 3.2%). The difference between the SMCI model and the SMC model is small in spite of the high and statistically significant negative error correlation (column 3). As it was discussed in section 1.3, a negative error correlation is expected for over-utilized treatments.

For the whole period, the estimated marginal probability of non-clinical factors was 3.2%. This means that 3.2% of healthy, non-risk women had a non-clinically required cesarean section in the period between 1999-2002, meaning that each year around 2,500 women have unnecessary c-sections in New Jersey. Even though the percentage of unnecessary c-sections is relatively small, a more detailed inspection of the results shows a positive trend in the doctor's incentive equation given by the year dummy variables. As a consequence, it is expected that most of the growth in the observed c-section rates in recent years can be explained by non-clinical factors rather than changes in health conditions in the population. To test this hypothesis, I compute the probability of appropriate c-sections measured as the marginal probability of c-sections due to health conditions only (see section 1.3 for further discussion). This estimated c-section rate is the rate without any non-clinical influence, and therefore without misclassification. Figure 1.2 compares both, the observed and the only-health-related estimated c-section rate. According to this figure, c-section rates in New Jersey grew from 22.5% in 1999 to 27.5% in 2002. However, the rapid growth started in 1997 after a long period in which cesarean rates were about 20%. The only-health-related estimated c-section rate is based on 1999-2000 data. For that period, it is shown that the rapid growth of c-sections was explained mainly by non-clinically required c-sections. Without physician incentives, the c-section rate in New Jersey would have remained almost constant at around 22%, i.e. just above the levels observed before 1997 when c-sections soared, and more in line with the recommended rate of 15% of Healthy People and the World Health Organization (1985).

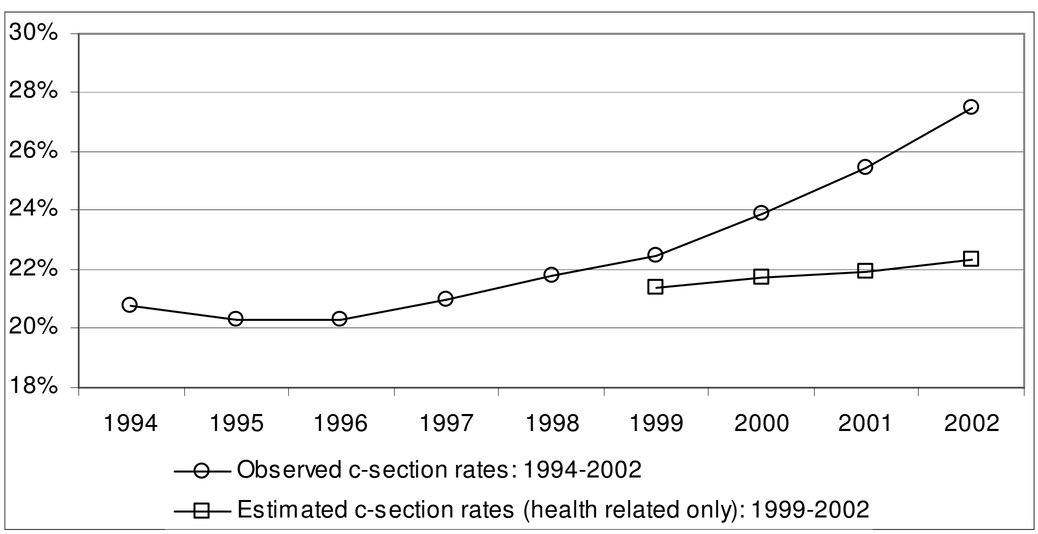

Figure 1.2: Cesarean Section Rates in New Jersey

What are the determinants of physician incentives? The estimated incentive equation is shown in Table A1.1. I structure the discussion of these results according to the two types of variables described before, and using marginal calculations for each estimated parameter. I found the expected direction for all variables related to signaling of the degree of patient-obtained medical information. Income and ethnicity are observed by physicians as indicators of patient observed medical information. For the average household income, an increment of five thousand dollars reduces the probability of c-section by 0.10%. This small but significant impact was also observed by Pauly (1980) for ambulatory care. Ethnicity had an important impact on the probability of c-sections. Black and Hispanic women have respectively 2.20% and 2.30% higher probability of having a c-section compared with other non-white ethnicities. White women have a 2.10% lower probability of c-section delivery. These results are consistent with previous literature (Aron et al. 2000).

Social support measured as married (or joint in life) women reduces physician incentives, implying a 1.90% lower probability of a c-section. Compliance to physician's influence was captured with women's employment status. As expected, full-time employed women have a 7% higher probability of c-section. Li et al. (2003) show also higher c-section rates for employed women.

The second type of variables related to factors that directly affect physician incentives also had the expected effect on c-sections. The most important and studied variable is payment source. With respect to non-HMO private insured patients, uninsured women are the least affected by doctor's incentives, with reduction in probability of c-section of around 9.20%. Medicaid beneficiaries are the next in lower influence with reduction in the probability of c-section of 3.40%. Finally, the capitated payment system of HMO reduces the probability of physician incentives, reducing probability of c-section in 1.30%. With respect to hospital size, the results confirm the supply-sensitive service hypothesis. In general, births in larger hospitals have higher probability of c-section. For a mid-size hospital, increasing births in 500 hundred per year raises the probability of a c-section in 0.10%. A similar argument is validated when I observe that women attended by more specialized physicians (Ob/Gym) have higher probabilities of c-section (2.60% more).

1.5 Conclusions

In this chapter I develop an econometric method to estimate over- and under-utilization of medical procedures. When a physician has incentives that keep him from choosing the appropriate treatment for a patient, the patient's health status loses correspondence with the observed treatment. This generates a problem whose characteristics and effects on estimation are analogous to a classification error. However, this particular measurement error is not random. This chapter proposes a structural model where the classification error is characterized by a physician behavior structure. That allows us to consistently estimate risk-adjusted utilization rates based on clinical factors only, and the probability of inappropriate treatments based on non-clinical factors (misclassification probability). Both measures can be neatly separated to test over- or under- healthcare utilization.

The results of the Monte Carlo study suggest that methodologies based on bivariate models (e.g. logit or probit) report biased estimates even when clinical or non-clinical factors are added as control variables. There are important interactions in the physician-patient relationship that can be captured by the structural misclassification model developed in this chapter. I apply the model to cesarean section deliveries performed in New Jersey from 1999 to 2002. The results show that around 3.2% of healthy, non-risky women had c-sections due to non-clinical factors. This rate implies that each year nearly 2,500 women have c-sections for non-medical reasons implying an excess cost of around $17.5 millions per year. Finally, it is estimated that non-clinical factors explain the rapid growth of c-section rates observed in New Jersey over these years.

CHAPTER 2: The Aftermath of the Peruvian Health Reform: Unnecessary C-Sections in the Private Health Care Sector

2.1 Introduction

The cesarean section rate in Peru as well as in many developing countries hides extreme problems. According to WHO statistics, 13% of Peruvian women had a c-section in 2000 which is in line with the WHO recommended c-section rate of 15%[3]. However, this national rate is highly heterogeneous among groups with different health care access. On the one hand, more than 40% of deliveries are not institutionalized, and are attended mainly in patient's home. For this group cesarean section is not accessible, so the c-section rate is almost zero, far below WHO recommendation. On the other hand, about 7% of deliveries are attended in private healthcare facilities. For this group the c-sections reach a rate of almost 50%, which is far above WHO recommendations.

In this chapter I focus on the high level of c-sections in the Peruvian private healthcare sector. Even though this problem has been observed before (SEPS, 2002; Alcázar y Andrade, 2000; Braschi, 2005), this is the first research that studies the problem at a national level, using an appropriate methodology to test over-treatment of cesarean deliveries. The study of c-sections in Peru gains higher importance due to its similarity with other Latin American countries. In a regional study, Villar et al. (2006) found that while the median rate of cesarean delivery in Latin America was 33%, in

[3] WHO (1985). This rate is based on c-section rates observed in countries with the lowest perinatal mortality rate.

private hospitals that rate reached 51%. Belizan et al. (1999) found similar results for the second half of nineties.

These high rates in Latin American's private hospitals have motivated studies focused on excess cesarean sections. The methodology used in these studies compares the crude observed c-section rate with the WHO's recommendation of 15% (Belizan et al., 1999; Hanvoravongchai et al., 2000). There are at least two limitations with that methodology. First, it is well known that c-sections are strongly and positively correlated to mother's age. Consequently, changes in demography toward elderly motherhood in recent years makes the more than 20 years old WHO recommendation of 15% obsolete. Using that guideline may over-estimate the excess number of cesarean sections. The second limitation is related to risk differences between mothers in the private and public sector. On average, age at motherhood is higher and number of children is lower in mothers in the private sector, which should increase the frequency of c-sections. Since those studies do not control for risk factors, the crude cesarean section rates in the private and public sector may produce misleading estimates of unnecessary c-sections. In this chapter I use the methodology presented in chapter 1 to estimate over-treatment which separates clinical and non-clinical determinants of cesarean sections.

The relevance of this study is twofold. First, over-use of cesarean sections has an important impact on mortality and morbidity of mother and newborn. In a study for Latin America, Villar et al. (2006) show that even after adjustment for risk factors, the rising rate of cesarean delivery was positively associated with postpartum antibiotic treatment and severe maternal morbidity and mortality. The rate of infection, the probability of subsequent placenta previa, and the risk of stillbirth are also higher in c-sections than in vaginal deliveries (NIH, 2006; Gray et al., 2007).

Cesarean surgery also increases the risk of newborn mortality and morbidity. In a recent study for the United States, MacDorman et al. (2006) found that neonatal mortality rates were 2.9 times higher among infants delivered by cesarean section than for those delivered vaginally. Similar results were found by Villar et al. (2006) using Latin American data. Studies have documented several possible effects of cesarean delivery on infant health, including respiratory morbidity, risk of persistent pulmonary hypertension, delayed neurologic adaptation, and delayed establishment of breastfeeding (Mac Dorman et al., 2006; NIH, 2006). In addition, Villar et al. (2006) found higher numbers of babies admitted to intensive care for seven days or longer even after adjustment for preterm delivery, and higher rates of preterm delivery.

Second, this study is relevant because it identifies the determinants of over-use of c-sections and estimates their effects. A better understanding of those factors will help to define policy recommendations oriented to reduce the problems in a more effective way.

This chapter is organized in five sections. The second section describes the health reform in Peru and explains how it has affected the private health care sector, creating incentives to over-treat. The third section describes the data and the methodology used. The fourth section estimates the model and reports results. Finally, conclusions for this chapter and policy recommendations are offered in the last section.

2.2 Health reform and cesarean sections in Peru

Under the "*Ley de Modernización de la Seguridad Social en Salud*" in 1997, Peru established the legal base to reform its health system. Following the experience of Chile and Colombia[4], health reform was based on two major changes: First, the creation of new managed healthcare organizations named *Entidades Prestadoras de Salud - EPS* that assumed the provision and management of health care services in the private sector. Second, the transformation of old agencies (the Ministry of Health, known as the MINSA, and the Health Social Security Service, known as ESSALUD) to assume new roles under the health reform.

The main goal of the reform was to reduce congestion in the Health Social Security Service, ESSALUD[5], by transferring demand to the private sector. This would allow the government to focus its efforts on improving access to health care in poor population. On the other hand, inducing the growth of the private health sector would increase competition, infrastructure investment, and health quality (Garcia, 2001.)

The health reform has increased the percentage of births attended in public health facilities, from 46% in 1994-1996 to 64% in 2002-2005 (see Table 2.1.) According to the Demographic and Health Surveys (DHS), the increment of institutionalized deliveries has focused on the public sector only through an important shift from home to public facilities. However, births in private facilities have stagnated at around 7% in the last two decades thus reflecting serious limitations of the reform in private health care provision.

---

[4] See Gonzales-Rossetti and Bossert (2000) for a comparative analysis of health reforms in Chile, Colombia and Mexico.

[5] Mainly in the case of normal health care (*capa simple*), that considers medical procedures with high frequency and low complexity.

Table 2.1: Births according to place of delivery: 1994-2005 (percentages)

Period	Public Sector	Private Sector	Home	Others	No Inform.	Total
1994-1996	46.1	7.6	44.8	1.5	0.0	100
1996-1999	51.9	6.7	40.2	1.1	0.1	100
2000-2002	54.6	6.9	36.7	1.8	0.0	100
2002-2005	64.2	7.1	27.4	1.3	0.0	100
Total	51.0	7.2	40.2	1.6	0.1	100

Source: DHS 1996, 2000 y 2004-2005

The success of the reform on the private sector market was questioned from its first years (Carbajal and Francke, 2000; García, 2001). Rather than expanding the private health market, EPS focus on health care provision to employees from medium and large companies who had previous access to alternative health care programs (health insurance, agreements with private hospitals, etc.) The limits to the expansion of private insurance were the extremely big informal sector, which includes more than half the total urban labor force, and the low salaries in smaller formal companies (Du Bois, 2005). As a consequence, affiliation to EPS was kept almost constant during the first years, and EPS could not reach optimal scale levels, thus generating profit losses. The number of managed healthcare organizations reduced from 4 to 2 EPS because of mergers and acquisitions. In general, the private health care market remains stagnant. Less than 3% of Peruvian population has some type of private health insurance, and out-of-pocket payment remains as the most important source of private health financing.

One important consequence of the health reform has been the change in the balance of power between private hospitals and EPS. The five largest private hospitals in Peru concentrate more than 45% of the private health market. Before the reform, individual employers seeking health coverage directly with private hospitals lacked power to negotiate prices or quality. After the reform, EPS grouped and organized employers' demand, thus gaining market power. The two EPS currently operating in the market[6] are financially related to important insurance companies that lead the health insurance market. Together, these two healthcare conglomerates now represent around 75% of private hospital's revenues. It was mainly through the exercise of market power in this highly concentrated market that privately managed healthcare organizations reached sustainability and consolidation seven years after the reform started.

EPS's higher negotiation power depresses health care prices, and together with stagnant market growth makes a tough combination for doctors' incomes. Under these circumstances, and under the umbrella of an unregulated market and non-existent disclosure of health outcomes, a fee-for-service physician payment system increases the incentives to overuse medical procedures. The informational advantage of the doctor with respect to the patient, allows them to induce health care demand without medical justification (McGuire, 2000). When private insurance enters the doctor-patient relationship, additional incentives are added (Cutler and Zeckhauser, 2000) because the price stops being a restriction on the patient's decision. In the doctor-patient-insurance relationship, it is easier for the physician to take actions to influence patient's preferences under the threat of higher medical risks. Given the difficulties in monitoring doctors' decisions in a private model, the only restriction for doctors is their professional ethics. If the financial incentives are high enough to encourage deviation from ethical standards, a doctor may potentially affect the exercise of patient's informed consent, and in the case of pregnant women, reproductive rights may be violated too.

There is substantial evidence of demand inducement in private health care provision (Gruber y Owings, 1996; Das, 2002). Overuse of medical procedures has been

[6] A new but small organization entered the market, starting operations in April 2007.

observed in Chile after the health reform (García, 2001; Murray, 2000). In Peru, the statistics from the EPS regulation authority (known as *Superintendence of EPS - SEPS*) show that the number of medical procedures has grown more than proportionally with enrollment. In a more detailed study for hypertension, asthma and respiratory infection, the SEPS (2002) shows overuse of diagnostic procedures and an excessive number of medical visits, reflecting a lack of standardization in the treatment of those diseases.

The overuse of cesarean sections for non-medical reasons in Peru has made news in the past (Braschi, 2005; Alcázar and Andrade, 2000; SEPS, 2002). However, this is the first national study that analyzes and quantifies the phenomenon from an economic standpoint. Peruvian national statistics show that the rate of c-sections in the private sector has almost doubled after the reform, reaching 48%, while in the public sector it remains around 19%. Figure 2.1 shows that the c-section rate in private facilities rose from 27% en 1991-1993 to 48% in 2002-2005.

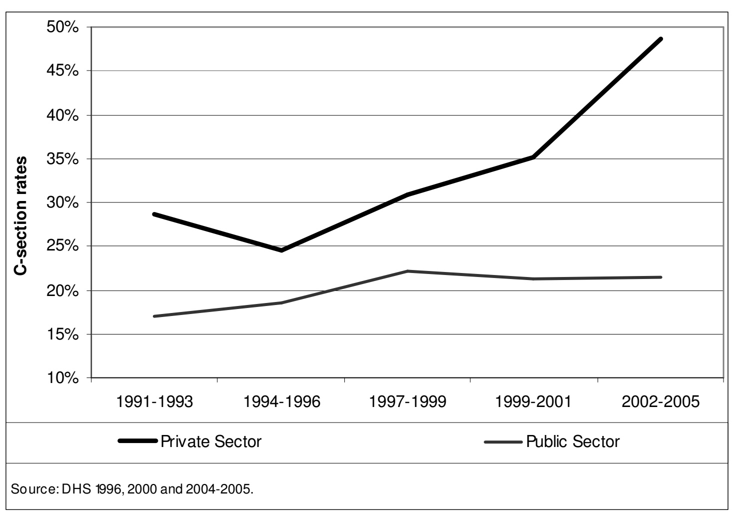

Figure 2.1: Trend of Cesarean Section Rates in Peru

Rapid growth in cesarean section rates has also been observed in Chile 10 years after the reform (García, 2001.) The Chilean data showed that 60 out of every 100 births were c-sections in the private sector, while in the public sector that number was only 25. The hospital ownership is an important variable explaining high rate of c-sections in many countries (Villar et al., 2006; Potter et al., 2001; Gomes et al., 1999; Belizan et al., 1999; Tussing and Wojtowycz, 1992; Gregory et al, 2001; Mossialos et al., 2005; Lin and Xirasagar, 2004). According to a recent study for eight countries in Latin America, Villar et al. (2006) found that the average c-section rate was 33%, but the rate in private hospitals was 51%.

When women have access to private health insurance the numbers become more dramatic. In the 2002-2005 period, the rate of cesarean surgeries in the private sector was 43% for women without insurance, and 66% for women with some private heath insurance. Again, a similar finding was observed in Chile ten years after the health reform (Murray, 2000.) Many studies have also shown that c-sections are more likely when women have insurance (Stewart-Hall, 2000; Mossialos et al., 2005; Tussing and Wojtowycz, 1992; Gomes et al., 1999; Aron et al., 2000; Luthy et al., 2003).

2.3 Estimation Methodology and Data

2.3.1 Methodology

To quantify the number of unnecessary c-sections I use the structural model of physician incentives developed in Chapter 1. The model captures the principal-agent problem in the patient-physician relationship (McGuire, 2000). It recognizes that health outcomes observed by the econometrician are the result of a decision process where the physician first observes patient health status and then chooses a clinically appropriate or inappropriate treatment based on her/his own monetary and non-monetary cost-benefit analysis, and the patient's attitude to his/her recommendation. The model is based on a game theory model of physician demand inducement. In the first stage patient health status (h) is determined by a set of observed (x) and non-observed (ε_h) clinical characteristics. The physician can observe patient health status defined by

$$h = x\beta + \varepsilon_h \qquad (2.1)$$

There are two possible treatments, natural delivery and cesarean section ($\tilde{y} = \{0,1\}$). The patient requires a c-section ($\tilde{y} = 1$) if her health status exceeds zero:

$$\tilde{y} = 1 \text{ if } h = x\beta + \varepsilon_h \geq 0$$
$$\tilde{y} = 0 \text{ otherwise}$$

The econometrician observes the treatment chosen by the doctor (y) but not the appropriate one ($\tilde{y}$). Without medical incentives to affect the appropriate treatment, $y = \tilde{y}$ and in that case any binary estimation model (i.e. probit or logit) would be consistent since the probability of observing the chosen treatment is equal to the probability of observing the appropriate treatment.

$$\Pr(y = 1) = \Pr(\tilde{y} = 1) = \Pr(h \geq 0)$$
$$\Pr(y = 0) = \Pr(\tilde{y} = 0) = \Pr(h < 0)$$

However, if physician opt by a unnecessary c-section, then the estimated binary model will produce inconsistent estimates since $\Pr(y=1) \neq \Pr(\tilde{y}=1)$. In those cases, the econometrician observes a treatment with "classification error".

The medical decision is modeled in a second stage. After observing patient health status, the physician chooses treatment based on net-incentives (i). The factors that characterize physician incentives reflect doctor utility after a monetary and non-monetary cost-benefit analysis that may include professional ethics, reputation, fear of law suits, leisure, income, etc. This stage can also be seen as the backward solution of a third stage in which the patient decides to accept (take the treatment) or reject (go to another doctor) the physician's recommendation. In this case, net-incentives (i) may include patient preferences and patient-obtained medical information. Net incentives depend on observed non-clinical variables (z) and non-observed variables (ε_i).

$$i = z\gamma + \varepsilon_i \tag{2.2}$$

Conditional on patient health status, the physician will choose a c-section if net-incentives are greater or equal to zero. However, the chosen treatment will be clinically inappropriate when patient health status is less than zero. In other words, c-section will be unnecessary for medical reasons with probability

$$\alpha_0 \equiv \Pr(y=1 \mid \tilde{y}=0) = \Pr(i \geq 0 \mid h < 0) \tag{2.3}$$

Consequently, the econometrician observes a c-section with probability:

$$\begin{aligned} \Pr(y=1) &= \Pr(h \geq 0) + \Pr(i \geq 0 \mid h < 0)\Pr(h < 0) \\ &= \alpha_0 + (1-\alpha_0)\Pr(h \geq 0) \end{aligned} \tag{2.4}$$

In Chapter 1 I present a generalization of this model to the case of under- and over- health care utilization, and its relationship with classification error models. The parameters β and γ of equations 2.1 and 2.2 can be estimated by maximum likelihood, maximizing the function:

$$L(\beta,\gamma) = \prod_{i=1}^{n} \Pr(y=1)^{y} \Pr(y=0)^{1-y} \tag{2.5}$$

I assume error terms $\varepsilon_h, \varepsilon_i$ are normally distributed and independent. The estimation was done using Gauss v5.0. Marginal effects are calculated as a one unit change for discrete variables, and a small change for continuous variables. In all cases I report the average of the change in probabilities keeping other variables at observed values.

2.3.2 Data

The data set used in this study is from the Demographic and Health Survey (DHS) implemented in Peru in the years 1996, 2000 and 2004-2005. The survey includes women in fertile age (15 to 49 years old) and children under 5 years old. This is a representative sample that allows inference at a national level. The DHS 1996 reports 15,639 births of liveborn babies between years 1991 and 1996. The DHS 2000 reports 12,222 births between years 1995 and 2000. The continuous DHS survey 2004-2005 reports 4,243 births between years 1999 and 2005. Data from non-institutionalized births was excluded to avoid misclassification error[7]. Therefore, this study focuses on institutionalized births within public or private health care facilities. The sample also considers only one birth for a mother, implying that in multiple gestations only the oldest baby is considered. Observations with missing information were also excluded. The final sample consists of 13,496 births of liveborn babies between 1991 and 2005.

Following the model, two sets of variables are defined. The first group describes clinical characteristics of pregnant woman (matrix x in equation 2.1). I follow the medical literature to identify key risk factors that can explain c-section for medical reasons (Keeler et al., 1997; Aron et al., 1998; DiGiuseppe et al., 2001; and Rahnama et al., 2006.) In this study, the following variables were identified: woman's age, multiple gestations, nulliparity, previous interrupted pregnancy, number of liveborn infants, fever during labor, convulsion during labor, size of newborn, number of pre-natal visits. Although the DHS survey does not have more information about medical conditions at delivery, the selected variables are found to be the most important variables predicting c-sections (Peaceman et al., 2002).

The second set of variables is related to physician net-incentives (matrix i in equation 2.2). Within this set two types of variables can be identified based on the model described before. The first type is related to direct monetary and non-monetary net-incentives. It includes: (i) place of delivery (public or private facility), (ii) access to private insurance, and (iii) date of delivery. The second type of variable is related to physician perception of patient information and preference. It includes (i) socio-economic characteristics, and (ii) attitude toward breastfeeding.

[7] Non-institutionalized births do not have c-section as an option. So even in the case of a medically required c-section, the observed outcome will be a vaginal delivery (misclassification), affecting predictivity of clinical risk factors in the estimation.

Hospital ownership has been shown relevant in different studies for Latin America (Villar et al., 2006; Potter et al., 2001; Gomes et al., 1999; Belizan et al., 1999), the United States (Tussing and Wojtowycz, 1992; Gregory et al, 2001), and other countries (Mossialos et al., 2005; Lin and Xirasagar, 2004). Financial incentives are greater in private health care facilities because physicians are paid on a fee-for-service basis; while in the public sector physicians are paid on a fixed salary basis.

Access to private health insurance is also important. In this case a moral hazard problem is added to the patient-physician relationship. Patients with insurance worry less about the price difference between one procedure and another. In particular, the higher cost of c-sections does not affect patient decisions as long as the insurance covers vaginal and cesarean delivery in the same way. Many studies have shown this mechanism showing that c-sections are more likely in insured women (Stewart-Hall, 2000; Mossialos et al., 2005; Tussing and Wojtowycz, 1992 and 1993; Gomes et al., 1999; Aron et al., 2000; Luthy et al., 2003). However, insurance companies may exert market power to control physician practices, and consequently to reduce the number of expensive treatments. A good example is c-section rates in the USA in the nineties. The increasing market power of HMO reduced the cost and number of treatments in hospitals (Kessler and McClellan, 2000; Gowrisankaran and Town, 2003), many times through strong penalties to physicians that deviate from the conservative cost practices imposed by HMOs (Robinson and Steiner, 1998). As a consequence, the probability of c-section falls when insurance is provided by an HMO (Das, 2002; Hueston and Sutton, 2000; Tussing and Wojtowycz, 1994). Unfortunately, there are two important caveats with the insurance variable in the DHS survey. First, DHS does not report insurance status at the moment of delivery, but at the moment of interview. The time difference between both events is at most 5 years. Therefore, the variable used in the study is a proxy with some measurement error. Second, DHS does not distinguish between types of private insurance, so it is not possible to identify health care organization (EPS) separately from other type of insurers. However, it is known that after the consolidation of the system, EPS have the largest share of the private health care market.

The study also includes the weekday effect of deliveries. Cesarean sections are convenient for the doctor because they can schedule births to maximize leisure time, and because they can set many births on the same day. Many studies (Gomes et al., 1999; Tussing and Wojtowycz, 1992) have shown that the probability of a c-section decreases on weekends and holidays, which is consistent with the hypothesis of c-sections by doctor's convenience. It is important to note that even though there are constraints on doctors in the public sector, doctors in ESSALUD have some degree of autonomy to schedule births (Alcázar and Andrade, 2000). It means that doctor's convenience may be present in the private and the public health sector too.

I include four variables related to physician perception of patient information and preference. First, it is expected that women from rural areas are more resistant to c-sections. Vaginal delivery is a traditional method of birth, and women in rural areas see c-sections as an intrusive and unnatural surgery. Cultural factors have been significant in many studies on c-sections. Aron et al. (2000), Hueston and Sutton (2000), and Gruber and Owings (1996) found that African American and Latin American women in the USA have lower probabilities of c-section than white women even after controlling for other socioeconomic factors.

It is expected that wealth is a factor in compliance to c-sections that in some degree is related to the *Too Posh To Push* movement against vaginal delivery. Wealth is measured by asset possession as calculated by the DHS program[8]. Women are grouped according to quintiles of household wealth. It is expected that wealthy women are more likely to have c-section based on a informed choice, as it has been observed by Leeb et al. (2005), Alves and Sheikh (2005), Potter et al. (2001), and Brugha and Pritze-Aliassime (2003).

Education is a variable that may reflect access to information. However, the low degree of public information related to c-sections and the risks they imply limit the effectiveness of this variable. Additionally, education by itself does not imply an attitude in favor of or against a c-section. In previous literature, it is shown that education indeed increases the probability of a c-section (Gomes et al., 2005; Patel et al., 2005; Tussing and Wojtowycz, 1992.)

Breastfeeding is an indicator of resistance to c-sections. It is well known that this procedure directly affects breastfeeding (Minkoff et al., 2004; Pérez-Escamilla et al., 1996,) therefore; a c-section would not be the choice of an informed woman who wants to breastfeed her newborn. Additionally, an informed woman who chooses a c-section for esthetical reasons would not see breastfeeding as an option. As a consequence, it is expected that a positive attitude toward breastfeeding will act as a factor against a c-section.

[8] The wealth index is estimated through a factor analysis procedure that takes into account a large number of household assets.

Table 2.2: Mean values of variables (in percentages)

Variables	Before EPS 1991 - 1999	After EPS 1999 - 2005
Delivery by c-section	20.2	22.9
Variables related to monetary and non-monetary incentives		
Birth was in a private facility	12.9	10.2
Birth in private sector and private insurance	2.1	2.1
Birth was Saturday	13.4	9.1
Birth was Sunday	11.8	8.8
Birth was non-working holiday	3.2	2.9
Variables related to perceived information and preferences of mother		
Woman grew up in a rural area	33.7	38.2
Woman with post-secondary education	28.0	28.5
Woman decided not to breastfeed	1.7	0.7
Household wealth: 1(lowest) to 5(highest)	3.5	3.4
Control variables: Clinical characteristics		
Age (in years)	27.5	27.5
Women is 35 years old or more	15.1	16.0
Multiple birth	0.8	1.2
Nulliparity	36.4	39.4
Number of children (in numbers)	2.5	2.3
Fever during labor	8.4	5.5
Convulsions during labor	4.3	2.8
Under-weight newborn	7.2	7.8
Over-weight newborn	7.4	5.8
Woman with history of interrupted pregnancy	20.2	17.7
Woman had 1 to 3 prenatal visits	12.1	6.4

Source: DHS 1996, 2000 y 2004-2005
Weighted values. Total number of un-weighted observations is 13,583

Table 2.2 shows all the variables used in the study, and their mean values in two periods: before (from September 1991 to June 1999) and after (from July 1999 to September 2005) the creation and operation of health care organizations (EPS) as reported by the EPS regulator (SEPS).

2.3.3 Results

Estimated parameters are presented in the Appendix (Table A2.1). Tables 2.3A-2.3C show the marginal effects of each group of variables described in the previous section. There are important results related to physician incentives (see Table 2.3A). First, there is a demand inducement problem in the private sector that has increased with the reform. The probability of a c-section when women go to a private facility has almost doubled after the reform. Today, this probability is almost 6% compared to 3.5% before the reform.

Second, private insurance adds incentives to choose a c-section, increasing the probability of c-section from around 2.4% before the reform to 4% after the reform. As was described in the previous section, insurance may have two opposite effects: First, insurance may increase c-section through a higher moral hazard effect. Since insurance reduces patient's monetary burden of c-section, it increases physician incentives to choose a c-section. That explains the positive marginal effect in c-section probabilities in both time periods. Second, insurance may decrease c-sections through higher market power and control on medical practices. That might explain why the marginal effect on c-sections has loose statistic significance after the health reform, a period where EPS started getting higher market power.

Third, doctor's convenience is an important factor explaining higher c-section rates after the reform due to the weekday effect. Before the reform, c-sections were less likely on Sundays (4.1%). After the reform, c-sections were even less likely on Sundays (7.5%) and now also less likely on Saturdays (3.8%). Non-working holidays were not statistically significant for the post-reform period. It is important to recall that doctor's convenience affects both the private and public health care sector.

Table 2.3A: Marginal Effects – Structural Misclassification Model Estimation

Variables related to monetary and non-monetary incentives		
	Before Reform	**After Reform**
Birth was in a private facility	0.035 *	0.059 *
Birth in private sector and private insurance	0.024 ***	0.040
Birth was Saturday	0.012	-0.038 **
Birth was Sunday	-0.041 **	-0.075 **
Birth was non-working holiday	-0.092 **	-0.019

Dependent variable is c-section birth. Before reform corresponds to period 09/1991-06/1999. After reform corresponds to period 07/1999-09/2005.
* significant at 10%; ** significant at 5%; *** significant at 1%

Figure 2.2 illustrates the effects of physician incentives. I use the results of the estimation to calculate the probability of having an unnecessary c-section given that delivery was in a public facility, private facility with insurance or private facility without insurance. A probability of zero indicates no unnecessary c-sections, and therefore no physician incentives to overuse c-sections. It can be seen that physicians in the private sector have increased incentives to perform c-sections after the reform, in contrast to the public sector where incentives have slightly reduced. Currently, a non-at-risk patient with private insurance has a probability of 41% to have a c-section, while if she has no insurance that probability would be 18%. It is important to notice that even though the reform has not affected incentives in the public sector, Figure 2.2 shows that around 5% of women have an unnecessary c-section in public health care facilities. This probability is related to non-monetary incentives (weekday effect), since monetary benefits of a c-section are limited for doctors in the public sector.

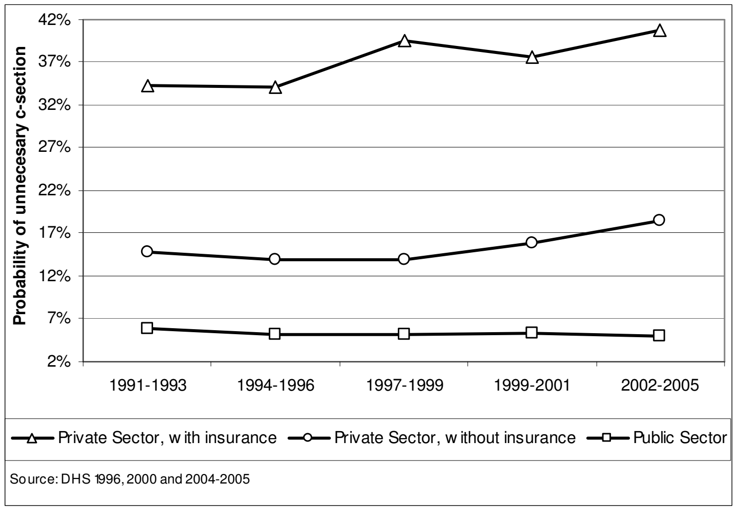

Figure 2.2: Physician Incentives-Probability of Unnecesary C-sections

Table 2.3B shows the marginal effects of the four variables related to physician perception of a woman's information and preference. As expected, women's origin plays an important role in resistance to cesarean sections due to cultural factors. This effect has grown after the reform. Before the reform, a woman who grew up in a rural area had 2% less probability of having a c-section, but after the reform the reduction in probability was more than 4%.

On the other hand, the effect of post-secondary education had a positive effect on c-sections. More education doesn't bring relevant information for a decision related to the method of delivery which shows the scarcity of public information about this topic. Another explanation may be related to preferences, which is consistent with the *Too Posh To Push* hypothesis described in the previous section. In general, women with post-secondary or higher education have a 2% higher probability of having a c-section.

Table 2.3B: Marginal Effects – Structural Misclassification Model Estimation

Variables related to perceived information and preferences of mother		
	Before Reform	**After Reform**
Woman grew up in a rural area	-0.020 **	-0.041 *
Woman with post-secondary education	0.022 *	0.021 **
Woman decided not to breastfeed	0.039 *	0.101 *
Household wealth: 1(lowest) to 5(highest)	0.056 *	0.045 *

Dependent variable is c-section birth. Before reform corresponds to period 09/1991-06/1999. After reform corresponds to period 07/1999-09/2005.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.3B also confirms the *Too Posh To Push* hypothesis but shows that this effect has not increased after the health reform. In general, wealthy women are more likely to have a c-section even after controlling for age, and access to private health care and private insurance. An increment in wealth (measured as a movement to the next wealth quintile) increases the probability of c-section by 4.5%. Before the reform, this increment was 5.6%, which indicates that preferences and attitude toward c-section changed after the reform.

The attitude towards breastfeeding turns to have an important effect on probability of c-section. This variable is related to a woman's preferences. Esthetic, comfort or work conditions may be behind both decisions. Women who choose not to breastfeed may also prefer a c-section because of the false myth that it is "fast, programmed and painless". The results show that before the reform, a negative attitude towards breastfeeding increased the probability of c-section by nearly 4%. However, after the reform, the increment in probability reached 10%. It is important to mention that even though the impact of non-breastfeeding on c-sections has increased, the percentage of

Peruvian women who do not breastfeed has slightly reduced in recent years to around 1.2%[9].

Finally, Table 2.3C shows the marginal effect of clinical factors on c-section. Even though this estimation can be used to calculate risk-adjusted c-section rates, the goal in this study is to analyze the impact of non-clinical factors on health outcomes. Therefore, clinical variables are used as controls. In general, the estimates are consistent with medical literature, and as should be the case, there is no significant difference between one period and another since estimates should be exogenous to the health system.

Table 2.3C: Marginal Effects – Structural Misclassification Model Estimation

Control variables: Clinical characteristics

	Before Reform	**After Reform**
Age (in years)	0.015 *	0.016 *
Women is 35 years old or more	0.066 *	0.102 **
Multiple birth	0.309 *	0.395 *
Nulliparity	0.076 *	0.113 **
Number of children (in numbers)	-0.049 *	-0.059 *
Fever during labor	0.101 *	0.010
Convulsions during labor	0.059 **	0.030
Under-weight newborn	0.116 *	0.180 *
Over-weight newborn	0.147 *	0.193 *
Woman with history of interrupted pregnancy	0.064 *	0.123 *
Woman had 1 to 3 prenatal visits	-0.250	0.060 **

Dependent variable is c-section birth. Before reform corresponds to period 09/1991-06/1999. After reform corresponds to period 07/1999-09/2005.
* significant at 10%; ** significant at 5%; *** significant at 1%

It is important to highlight the case of prenatal visits. Before the reform this variable was not significant. However, after the reform the variable becomes significant but with a positive sign, indicating a higher probability of c-section. This might be related

[9] Information from the DHS 1996, 2000 and 2004-2005.

to changes in standards of prenatal visits in the public and private sector. In both cases, the average number of visits increased from 1 to around 8 visits[10]. This may reflect an institutional effort to increase quantity but not quality of prenatal visits, which may reduce their predictive power. This argument is consistent with Guzman (2002), who suggests that the poor predictive power of prenatal visits in Peru implies that the focus of the program is not on quality but on production where “numbers are more important”.

The methodology used in this study allows separating out the impact of non-clinical factors on health outcomes. By removing the effect of non-clinical factors, it is possible to come up with a clinically appropriate c-section rate that is only based on medical factors. Figure 2.3 shows the trend of observed c-section rates and the clinically appropriate c-section rate, which can be viewed as a medical guideline for c-sections. It is clear that the gap between both rates widened after the health reform. For the period 2002-2005, the observed rate was 48.5% while the clinically required c-section rate should have been around 17.8%.

[10] Information from the DHS 1996, 2000 and 2004-2005.

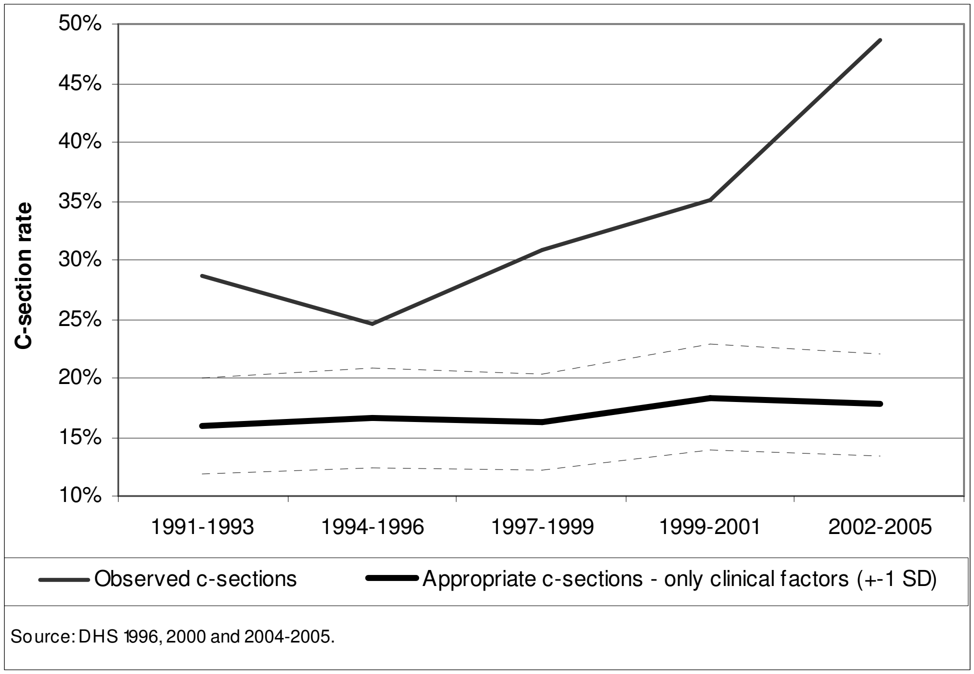

Figure 2.3: Private Sector – Observed C-Sections and C-sections without Physician Incentives

It is important to notice that the clinically required c-section rates estimated in the model are consistent with WHO recommendations. The WHO recommends a rate of 15%, but as expected, the demographic difference between public and private sector patients (motherhood at a later age in the case of women in the private sector) may imply that this rate should be higher for patients in the private sector. This study produces a guideline for c-sections that implies an average appropriate c-section rate of near 18%. This medical guideline has a slightly positive trend that is consistent with demographic changes toward mature-aged motherhood.

Based on the estimated clinically required c-section rates, and using the post-reform results presented in the Appendix, Table 2.4 reports the estimated number of women who have had a cesarean section without medical justification. It is estimated that each year more than 13 thousand Peruvian women had an unnecessary c-section. According to information on mean costs of c-sections and vaginal deliveries reported by the SEPS, these unnecessary c-sections represent more than 6.7 million dollars per year paid in excess. This excess cost is paid mainly by insurers (EPS, and other insurance companies) and households. Indirect costs related to higher frequency of readmission to hospital, medical visits and drugs related to c-sections are not included in this cost.

Table 2.4: Number of Unnecessary C-sections and Excess Cost

Description [1/]	Number of women
Total number of births in private facilities	42,830
Current number of women with c-sections	20,804
Number of women with c-sections if doctor's financial incentives are eliminated [2/]	7,615
Number of women with unnecessary c-sections	13,189
Excess Cost of unnecessary c-sections [3/]	US$ 6,727,174

[1/] Projections use means weighted by sample weights. Means are for years 2000-2005. A yearly average of 628 thousand births was considered (INEI, 2001.)
[2/] Based on the estimation of the structural misclassification model post-reform (see Appendix)
[3/] Based on EPS mean costs of c-sections and vaginal deliveries reported by the SEPS (fourth quarter 2005)

2.4 Conclusions and Recommendations

This chapter shows that the reform of the Peruvian private health system has increased doctor's financial incentives to overuse medical procedures. After the reform, managed healthcare organizations got enough market power to push healthcare prices down, but were unable to increase the access to private health. With lower incomes and a stagnant market, doctors have higher incentives to increase cesarean sections because they are more profitable than vaginal deliveries. This explains why, after the reform, the c-section rate in the private sector has grown from 27% to 48%, but this rate reaches 66%

when women have access to private insurance. In the public sector, however, the c-section rates remain in around 19%. This study shows that each year more than 13 thousand Peruvian women have c-sections without medical justification in the private sector. It represents more than 6.7 million dollars per year paid in excess.

Even though this study is concerned with the Peruvian case, reforms in the private health systems of many countries in Latin America have had an important impact on the dramatic increase of cesarean sections. In a very detailed document about health reforms in Latin America, Mesa-Lago (2005) found strong market concentration and barriers to competence among managed healthcare organization in many countries of the region. The case of Chile, Colombia, Argentina and Mexico are highlighted, and precisely those countries show high rates of c-sections. The present hypothesis, now confirmed for Peru, requires to be studied for the whole region.

Beyond the economic causes and consequences on health expenditure, the high rate of cesarean sections in Latin America has opened a new threat on women's empowerment. After a decade of reforms in their health systems, most Latin American countries have left vulnerable patients' rights within the private health care market. The right to choose the mode of delivery is implicit in the definition of reproductive rights stated by most international organizations. The concept of safe motherhood (ICPD, 1994) includes the "adequate delivery assistance that avoids excessive recourse to cesarean sections". A cesarean section for non-medical reasons and without a woman's informed consent constitutes a fault that transgresses medical professional ethics, as is recognized nowadays by many obstetric and gynecology associations[11]. In a recent statement[12], the

[11] See for example NIH (2006), Christilaw (2006), Minkoff et al. (2004). Meredith (2005).
[12] FIGO Statement on Cesarean Section. January 2007. http://www.figo.org/Cesarean.asp

International Federation of Gynecology and Obstetrics (FIGO) considers the c-section procedure as a surgery for safe maternity care that should be undertaken only when indicated to enhance the well-being of mothers and babies and improve outcomes. FIGO considers c-sections for non-medical reasons "*to fall outside the bounds of best professional practice*".

Medical literature has shown a cesarean section is more risky than a vaginal delivery. However, in contrast with any other surgery for non-medical reasons (e.g. plastic surgery), a c-section also poses risks for the fetus. As a cornerstone principle of empowerment, women must have the freedom and autonomy to make an informed choice that considers the risk for themselves and for their fetuses. In that regard, the doctor's position should be to bring accurate and understandable information about the modes of delivery, highlighting that c-sections must be carried out only for specific medical reasons. Therefore, c-sections must be the result of a truly informed consent from women.

There is an important inertial component in the number of cesareans. On one hand, the influence of doctor's inducement and the large number of c-sections are changing women's preferences to favor cesarean surgeries, thus threatening women's empowerment. On the other hand, women who had a c-section are most likely to have c-sections in future pregnancies. Unnecessary cesarean sections not only destroy basic social rights, but also create inefficiencies in resource allocation. These excess resources may be used more efficiently in improving healthcare quality and access for the poor.

In order to establish solutions oriented to reduce the number of unnecessary c-sections, it is important to break doctor's incentives and woman's informational disadvantages. Because this is a problem with an economic origin, this study recommends four main economic policies to fight against cesarean surgeries that violate women's rights. The first two recommendations are oriented to reduce doctor's incentives. The last two recommendations are oriented to reduce women's informational disadvantage.

First, doctor's incentives may be reduced with monitoring. Managed healthcare organizations, hospitals and health regulators may participate in hospital audit committees to review cesarean surgery decisions or give a second opinion for programmed c-sections. Salinas et al. (2004) studied the experience of the *Hospital Clínico de la Universidad de Chile*. In that hospital, a process of medical auditing was established for obstetric services to analyze and control systematically some obstetric practices. With the implementation of this audit, the c-section rate fell from 44.9% to 37.1%. A similar program was also implemented in the Peruvian health social security service (ESSALUD) through an auditory program that reviews clinical files and reasons for c-sections. With respect to second-opinion programs, Sloan el al. (2000) report the case of the major maternity hospital in Quito, Ecuador (*Maternidad Isidro Ayora*). A program to provide patient co-management for cesarean section candidates was implemented in the hospital, requiring a second opinion for all non-mandatory cesarean candidates. As a result, cesarean section rates fell from 26.6% to 22.1% in only two months. In the USA, Robinson and Steiner (1998) describe how healthcare management organizations (in particular Health Maintenance Organizations, HMO), use their market power to impose high penalties -directly or indirectly- on physicians that consistently violate the cost conservative practices imposed by HMO. Monitoring is an important tool to increase the cost of doctor's inducement. If professional ethics are not enough to avoid unnecessary c-sections, monitoring has proven to be an effective tool.

Second, doctor's incentives may be reduced by eliminating the higher profitability of c-sections. A cesarean surgery is more attractive for physicians than a vaginal delivery because it pays more and it allows the doctor to program many on a same day. Therefore, if the price difference between modes of deliveries is eliminated, then c-section rates should fall considerably. This has been observed in the USA where the growth of HMOs was accompanied by a reduction of c-section rates. Das (2002) highlights that an important factor behind this reduction is the capitation of physician fees that, in the context of deliveries, implies that managed care reimburses cesareans and vaginal deliveries equally. In that regard, managed healthcare organizations in Latin America may structure a price scheme similar to HMOs that reduces or eliminates direct pecuniary benefits of c-sections.

Third, women's informational disadvantage may be reduced with higher disclosure of information. An important concern in the reform of the health systems in Latin America is the weak regulation of the private health system (Mesa-Lago, 2005). A global and active regulator of the health system is required to guarantee more and better public information about cesarean rates. The experience of the USA is useful in this regard, where many states (e.g. Virginia, New Jersey, Pennsylvania, among others) have started to publish risk-adjusted cesarean rates at the level of hospitals and even doctors. Risk-adjusted cesarean rates give a fair indication of cesarean surgeries for non-medical reasons, and they are useful to rank hospitals or doctors who potentially overuse medical procedures. The health regulator should get enough information to compute these indicators, and to publicly disclose a ranking of hospitals and physicians with the highest rates of risk-adjusted c-sections. This information will be valuable for women in their informed choice process.

Fourth, women's informational disadvantage may be reduced with informed public opinion. The government, as well as all organizations that promote women's reproductive rights, must have an active position to organize programs oriented to inform the public about the risks of cesarean section and true medical reasons that justify a c-section. Public opinion must be alerted to the threat of high rates of cesareans on women's reproductive rights and higher risks for the mother and the fetus. In that regard, public opinion campaigns like the World Respected Childbirth Week are highly encouraged to revert changes in preferences and to generate higher resistance to doctor's inducement. Public relations efforts to disseminate information about the problem to key opinion leaders are also relevant.

This study has shown that the problem of unnecessary cesarean section in the private healthcare market in Peru has an economic origin with dangerous consequences for women's reproductive rights. The reform in the private health system has increased c-section rates without offering adequate regulation to control the problem. It is important to establish public policies directed to break the origin of the problem that rests in doctor's incentives and women's informational disadvantages. The solution must consider the participation of affected actors and must be oriented to inform the publis of the threat of this phenomenon to women's reproductive rights. It is urgent to take action to reverse the current trend of cesarean sections in Latin America and to stop what looks today a truly regional epidemic in private healthcare provision.

CHAPTER 3: A Semiparametric Estimation of a Structural Misclassification Model

3.1 Introduction

In Chapter 1 a structural misclassification model based on a two stage game is developed to account for non-clinical factors and its effect on health status variables. The structural misclassification model considers a physician that observes patient health status in the first stage, and decides to do the appropriate or inappropriate treatment based on incentives and observed patient preferences and medical information in the second stage. The model is able to separate out the impact of non-clinical factors on health outcomes, allowing over-treatment, defined as procedures that cannot be justified by health status variables, to be identified. However, this model heavily relies on parametric assumptions about the distribution of unobserved error terms.

The goal of this chapter is twofold. First, it explores how sensitive the estimates of the structural misclassification model are to distributional assumptions. Based on Monte Carlo simulations I evaluate the robustness of the structural model to misspecified error distribution. The second goal is to analyze the impact of reducing the parametric assumptions of the structural misclassification model by using a double-index semi-parametric estimation (Klein and Vella, 2008; Klein and Shen, 2008). The results of both models, the parametric and the semi-parametric model, are compared using data on cesarean sections in New Jersey.

The second section of this chapter describes the structural misclassification model in its parametric version and extends it to a semi-parametric form. The third section discusses the results of a Monte Carlo simulation that evaluates the robustness of the parametric and semi-parametric methods. The fourth section provides an application to the case of cesarean section deliveries in New Jersey, and compares the results of both models. The last section provides conclusions to this chapter.

3.2 Semi-parametric estimation of a Structural Misclassification Model

The structural misclassification model captures the principal-agent problem in the patient-physician relationship (McGuire, 2000). It recognizes that health outcomes observed by the econometrician are the result of a decision process where the physician first observes patient health status and then chooses a clinically appropriate or inappropriate treatment based on her/his own monetary and non-monetary cost-benefit analysis and the patient's attitude to his/her recommendation. The model is based on a model of physician demand inducement. In the first stage patient health status (H) is determined by a set of observed (X) and non-observed (ε_H) clinical characteristics. The physician can observe patient health status defined by the index

$$H = x\beta + \varepsilon_H \tag{3.1}$$

This is the health status equation that describes the positive relationship between health status and the index H, so a higher H implies worse health. There are two possible treatments: $\tilde{y} = \{0,1\}$. Patient requires treatment $\tilde{y} = 1$ if her health status exceeds zero:

$$\tilde{y} = 1 \text{ if } H = x\beta + \varepsilon_H \geq 0$$
$$\tilde{y} = 0 \text{ otherwise}$$

The econometrician observes the treatment chosen by the doctor (y) but not the appropriate treatment ($\tilde{y}$). Without incentives to alter the appropriate treatment, the physician chooses $y = \tilde{y}$. However, if physician chooses an inappropriate treatment, then $\Pr(y=1) \neq \Pr(\tilde{y}=1)$, and any binary model will produce inconsistent estimates. In those cases, the econometrician observes a treatment with "classification error".

The medical decision is modeled in a second stage. After observing patient health status, the physician may choose an inappropriate treatment based on net-incentives (I). The factors that characterize physician incentives reflect doctor utility after a monetary and non-monetary cost-benefit analysis that may include professional ethics, reputation, fear of suits, leisure, income, etc. This stage can also be seen as the backward solution of a third stage in which patient decides to accept or reject the physician's recommendation. In this case, net incentives (I) may include patient preferences and patient obtained medical information. Net incentives depend on observed non-clinical variables (z) and non-observed variables (ε_I), and are defined by the index

$$I = z\gamma + \varepsilon_I \tag{3.2}$$

Conditional on patient health status, the physician will choose a treatment if net-incentives are greater than or equal to zero. The chosen treatment will be clinically inappropriate when patient health status is considered at risk. In other words, over-treatment occurs with probability

$$\alpha_0 \equiv \Pr(y=1 \mid \tilde{y}=0) = \Pr(I \geq 0 \mid H < 0) \tag{3.3}$$

The degree of over-treatment can be measured by α_0, and a test of over-treatment can be set by evaluating the null hypothesis $H_o : \alpha_0 = 0$. For this chapter, I impose the restriction that under-treatment is not present. Notice that the econometrician observes the treatment $y = 1$ and $y = 0$ with probabilities:

$$
\begin{aligned}
\Pr(y = 1) &= \Pr(H \geq 0) + \Pr(I \geq 0 \mid H < 0)\Pr(H < 0) \\
&= \alpha_0 + (1 - \alpha_0)\Pr(H \geq 0) \\
\Pr(y = 0) &= \Pr(I < 0 \mid H < 0)\Pr(H < 0) \\
&= (1 - \alpha_0)(1 - \Pr(H \geq 0))
\end{aligned}
\tag{3.4}
$$

Maximum likelihood estimation can be implemented by employing these probabilities to construct the likelihood. See Chapter 1 for a generalization of this model to the case of under- and over- health care utilization, and its relationship with classification error models. The model in equation 3.4 can be estimated parametrically using normal or logistic probability distributions, yielding a bivariate probit or logit respectively. Notice that because only two of four possible outcomes are observed, this becomes a partial observability model. In particular, the model in equation 3.4 is identical to the one studied by Poirier (1980) under the assumption of normality.

For the binary model, Klein and Spady (1993) developed a single-index semi-parametric maximum likelihood (SML) estimator based on regular kernel estimation. The structural misclassification model implies the estimation of bivariate probabilities of the form:

$$
\begin{aligned}
P_0 &\equiv \Pr(y = 0) = \Pr(I < 0, H < 0) = F(-z\gamma_0, -x\beta_0) \\
P_1 &\equiv \Pr(y = 1) = 1 - P_0
\end{aligned}
\tag{3.5}
$$

Therefore, the probabilities are functions of two indexes: $z\gamma_0, x\beta_0$. In a parametric framework, the functional form for F is assumed, while in a semi-parametric framework the function F is left unspecified. The generalization of SML estimation from single-index to multiple indices is straightforward. However, root-n normality, which is desirable for making inferences and testing hypothesis, is not guaranteed with regular kernels. An alternative to recover root-n normality is to use higher-order kernels. Ichimura and Lee (1993) explored this alternative in the context of least square estimation by extending the semi-parametric least square (SLS) estimation of single-index models (Ichimura, 1993) to multiple indices. For binary response models, Lee (1995) extended Klein and Spady's SML model to a SML multiple-index model using higher-order kernels with a penalty function in the likelihood function to get rid of negative values generated by higher-order kernels. Although higher-order kernels are theoretically effective to reduce bias and reach root-n normality, their performance in finite samples is often less than desirable. A second alternative to get root-n normality is the one proposed by Klein and Shen (2008) who use a bias correction technique to reduce the bias in single-index SLS estimation. This technique uses a two-stage estimation based on regular kernels and a final step for bias correction. Klein and Vella (2008) extend the methodology to a double-index binary model using SML. In this chapter I use the results of Klein and Vella for my double-index binary model with partial observability.

To guarantee identification, I assume that the sets of variables x and z in equations 3.1 and 3.2 each include at least one continuous variable: x_1 and z_1 respectively. The indexes are then normalized up to scale and intercept to be:

$$v_H = x_1 + x_2\tilde{\beta}_0$$
$$v_I = z_1 + z_2\tilde{\gamma}_0$$

Let $\theta \equiv \{\tilde{\beta}_0, \tilde{\gamma}_0\}$, and Θ the parameter space of θ. Given a random sample of size n, P_0 can be estimated semi-parametrically using a kernel regression function:

$$P_{n,0}(x_{2j}, z_{2j};\theta) = \frac{A_n(y \mid x_{2j}, z_{2j};\theta)}{A_n(1 \mid x_{2j}, z_{2j};\theta)}, \tag{3.6}$$

Where

$$A_n(d \mid x_{2j}, z_{2j};\theta) = \frac{1}{n-1}\sum_{k \neq j} d_k \frac{1}{h_1} K\left(\frac{v_{I,j} - v_{I,k}}{h_1}\right)\frac{1}{h_2} K\left(\frac{v_{H,j} - v_{H,k}}{h_2}\right)$$

Here, y is the observed treatment choice, K is a regular kernel, i.e. any density function symmetric around zero, and h_1 and h_2 are bandwidth parameters. In this chapter I consider the following SML estimation:

$$\max_{\theta \in \Theta} \frac{1}{n}\sum_{j=1}^{n} \tau_j \left[(1 - y_j)\ln P_{n,0}(x_{2j}, z_{2j};\theta) + y_j \ln\left(1 - P_{n,0}(x_{2j}, z_{2j};\theta)\right)\right] \tag{3.7}$$

Where τ_j is a trimming indicator that trims out small values of the estimated density function $A_n(1 \mid x_{2j}, z_{1j};\theta)$, to guarantee well defined estimated probabilities. Klein and Shen define the trimming function τ_j over the set of indexes, and therefore two-stage estimation is required. In the first stage the estimation is done using a trimming over the set of regressors, while in the second stage estimated indexes from the previous stage are used to define the new trimming function. Finally, after the two-stage estimation, a bias correction is done without implying further optimization.

3.3 Robustness to misspecified error distribution: A Monte Carlo study

In this section I evaluate the robustness of the parametric and semi-parametric estimators to misspecification of the error distribution. The aim of this section is two fold. First, I want to quantify the loss of consistency in the parametric model when the normality assumption is violated. Second, I want to evaluate the robustness of semi-parametric estimation when normality is questionable. Two Monte Carlo designs are simulated. In the first design the data is drawn based on two correlated standard normal errors. In the second design the data is drawn based on two correlated standardized non-central t-distributed error with 4 degrees of freedom and non-centrality parameter 3, which produces a highly skewed distribution. For both designs the true model representing the health status equation 3.1 and the incentive equation 3.2 is:

$$\begin{aligned} H &= -4.0 + 0.5x_1 - x_2 + 2x_3 + \varepsilon_H \\ I &= -3.5 - 1.0z_1 + 1.5z_2 + z_3 + \varepsilon_I \end{aligned} \tag{3.8}$$

Covariates x and z include dummy variables and continuous variables that were drawn from uniform distributions and trimmed chi-squared distributions to avoid outliers. To be consistent with actual data, it is assumed that over-treatment, which is represented as misclassification in the model, is small, and therefore the probability associated to the incentive equation is set to be around 10% while the probability associated to the health status equation is around 50%. Although only the combination of the two probabilities described by equation 3.4 is observed, the implication is that information from the incentive equation is lower than from the health status equation. The error terms $\varepsilon_h, \varepsilon_i$ are jointly distributed with correlation $\rho = 0.15$. Both designs of the Monte Carlo study consider 500 independent random draws. To evaluate the impact of sample size, I consider samples of 1000 and 2000 observations. Following Chapter 1, the parametric case is estimated by ML assuming normal distribution and correlated errors, which yield a bivariate probit with partial observability. The semi-parametric ML case is estimated using bias correction as in Klein and Vella (2008). As described in the previous section, this semi-parametric method produces root-n consistent and asymptotically normal estimators without making distributional assumptions on the unobserved errors.

Tables 3.1A and 3.1B report the sample mean, standard error and root MSE of parameters estimated using a sample size of 1000 observations for the case when errors are normally and non-central t- distributed respectively. Under normal errors (Table 3.1A), the parametric model produces consistent coefficients for both the health status and the incentive equation. As is reasonable, the MSE is higher for discrete variables because of their lower variability.

Table 3.1A: Monte Carlo Simulation - Error Terms have a Normal Distribution, sample size 1000

True Parameters		Parametric Model Estimates		2-Stage BC Semi-parametric Model Estimates
Coefficient	Ratio	Coefficient	Ratio	Ratio
β_0 = -4.00		-4.05 *(0.338)* *[0.342]*		
β_1 = 0.50		0.51 *(0.083)* *[0.083]*		
β_2 = -1.00	β_2/β_1 = -2.00	-1.02 *(0.177)* *[0.178]*	-2.06 *(0.471)* *[0.475]*	-2.08 *(0.515)* *[0.52]*
β_3 = 2.00	β_3/β_1 = 4.00	2.02 *(0.154)* *[0.155]*	4.09 *(0.645)* *[0.651]*	4.19 *(0.709)* *[0.734]*
γ_0 = -3.50		-3.76 *(0.966)* *[1.001]*		
γ_1 = -1.00		-1.02 *(0.189)* *[0.19]*		
γ_2 = 1.50	γ_2/γ_1 = -1.50	1.73 *(0.929)* *[0.957]*	-1.75 *(1.016)* *[1.047]*	-1.48 *(0.448)* *[0.449]*
γ_3 = 1.00	γ_3/γ_1 = -1.00	1.02 *(0.126)* *[0.127]*	-1.03 *(0.2)* *[0.201]*	-1.06 *(0.229)* *[0.236]*
ρ = 0.15		0.14 *(0.393)* *[0.393]*		

N=1000, 500 simulations. Error terms are jointly standard normal distributed.
Standard errors in parenthesis and root MSE in brackets.

Table 3.1B: Monte Carlo Simulation - Error Terms have a Non Central t-Distribution, sample size 1000

True Parameters		Parametric Model Estimates		2-Stage BC Semi-parametric Model Estimates
Coefficient	Ratio	Coefficient	Ratio	Ratio
β_0 =-4.00		-5.75 *(1.114)* *[2.074]*		
$\beta_1 = 0.50$		0.72 *(0.152)* *[0.264]*		
β_2 = -1.00	β_2/β_1 = -2.00	-1.44 *(0.312)* *[0.539]*	-2.05 *(0.392)* *[0.395]*	-2.08 *(0.409)* *[0.417]*
$\beta_3 = 2.00$	$\beta_3/\beta_1 = 4.00$	2.83 *(0.491)* *[0.96]*	4.01 *(0.526)* *[0.526]*	4.19 *(0.557)* *[0.589]*
γ_0 = -3.50		-3.00 *(1.484)* *[1.565]*		
γ_1 = -1.00		-0.77 *(0.274)* *[0.36]*		
$\gamma_2 = 1.50$	γ_2/γ_1 = -1.50	1.21 *(1.416)* *[1.446]*	-1.62 *(1.336)* *[1.342]*	-1.40 *(0.664)* *[0.672]*
$\gamma_3 = 1.00$	γ_3/γ_1 = -1.00	0.81 *(0.171)* *[0.253]*	-1.16 *(0.358)* *[0.391]*	-1.14 *(0.354)* *[0.382]*
$\rho = 0.15$		0.34 *(0.583)* *[0.613]*		

N=1000, 500 simulations. Error terms are jointly standard non-central t-distributed with 4 degrees of freedom and non-centrality parameter 3.
Standard errors in parenthesis and root MSE in brackets.

In the parametric model I also compute the ratio to the first slope to make it comparable to the semi-parametric estimates. It is important to highlight that parametric probabilities and marginal effects do not depend on parameter ratios but on coefficients.

As it is described in equation 3.5, the probability depends on the specific functional form assumed and the true coefficients. If parameter estimates $\hat{\gamma}$ and $\hat{\beta}$ are close to the true values γ_0 and β_0, and if the assumed functional form F is correct, then the estimated probability function $\hat{P}_1$ will be close to the true probability function P_1. As a result, marginal effects will also be accurate, with such effects measuring probability changes that come from changes in individual exogenous variables. Therefore, it is important and informative to compare not only parameter ratios, but also probabilities and marginal effects under parametric and semi-parametric methods

For the semi-parametric case, it is instructive to note non functional form is assumed, so true probabilities are written as the conditional expectation of the binary dependent variable. Namely:

$$
\begin{aligned}
P_1 &= E[y \mid z\gamma_0, x\beta_0] = E[y \mid \gamma_{0,0} + z_1\gamma_{0,1} + z_2\gamma_{0,2} + ..., \beta_0 + x_1\beta_{0,1} + x_2\beta_{0,2} + ...] \\
&= E[y \mid z_1 + z_2\frac{\gamma_{0,2}}{\gamma_{0,1}} + ..., x_1 + x_2\frac{\beta_{0,2}}{\beta_{0,1}} + ...] \\
&= E[y \mid z_1 + z_2\tilde{\gamma}_0, x_1 + x_2\tilde{\beta}_0]
\end{aligned}
$$

This is the basis for equation 3.6. In writing the probability in this form, note that the expectation remains the same whether I condition on both indices ($z\gamma_0$ and $x\beta_0$) or whether linear transformations of the indices are taken. For this reason, coefficients are normalized up to constant and scale, and in this case the normalization is done through the ratio to the first slope.

Table 3.1A compares parameter ratios for the two methods for the case in which the error distribution is specified correctly. From this table, the semi-parametric estimation performs almost as well as the parametric estimation according to the MSE, and for some variables in the incentive equation it performs better.

Table 3.1B reports the results under the assumption of errors following a non-central t-distribution highly skewed to the left. The parametric model, which now erroneously assumes normality, produces inconsistent estimates with very large MSE for their coefficient estimates. The MSE of the ratios derived from the parametric coefficient estimates are similar to those reported in Table 3.1A under normality. Although these ratios are close to the truth, the misspecification error reduces the capability of the model for inference and prediction since those results are based on coefficient estimates.

To evaluate the impact of an increment in the sample size, I reproduce the same tables with 2000 observations. Tables 3.2A and 3.2B report these results. Two important conclusions can be obtained. First, in the parametric model, increasing the sample size reduces the MSE mainly through a lower standard error. The bias resulting from erroneously assuming non-normality does not disappear with a larger sample size. In Table 3.1B and Table 3.2B the biases of coefficient estimates are the same, and even bigger in the incentive equation estimated with a larger sample size (Table 3.2B).

Table 3.2A: Monte Carlo Simulation - Error Terms have a Normal Distribution, sample size 2000

True Parameters		Parametric Model Estimates		2-Stage BC Semi-parametric Model Estimates
Coefficient	Ratio	Coefficient	Ratio	Ratio
β_0 =-4.00		-4.03 *(0.23)* *[0.232]*		
β_1 =0.50		0.50 *(0.059)* *[0.059]*		
β_2 =-1.00	β_2/β_1 =-2.00	-1.00 *(0.128)* *[0.128]*	-2.02 *(0.329)* *[0.329]*	-2.03 *(0.358)* *[0.359]*
β_3 =2.00	β_3/β_1 =4.00	2.01 *(0.105)* *[0.106]*	4.03 *(0.458)* *[0.459]*	4.09 *(0.502)* *[0.511]*
γ_0 =-3.50		-3.55 *(0.414)* *[0.417]*		
γ_1 =-1.00		-1.02 *(0.137)* *[0.139]*		
γ_2 =1.50	γ_2/γ_1 =-1.50	1.54 *(0.373)* *[0.374]*	-1.53 *(0.394)* *[0.394]*	-1.45 *(0.343)* *[0.347]*
γ_3 =1.00	γ_3/γ_1 =-1.00	1.01 *(0.089)* *[0.089]*	-1.00 *(0.145)* *[0.145]*	-1.03 *(0.166)* *[0.17]*
ρ =0.15		0.14 *(0.266)* *[0.266]*		

N=2000, 500 simulations. Error terms are jointly standard normal distributed.
Standard errors in parenthesis and root MSE in brackets.

Table 3.2B: Monte Carlo Simulation - Error Terms have a Non Central t-Distribution, sample size 2000

True Parameters		Parametric Model Estimates		2-Stage BC Semi-parametric Model Estimates
Coefficient	Ratio	Coefficient	Ratio	Ratio
β_0 = -4.00		-5.65 *(0.856)* *[1.854]*		
β_1 = 0.50		0.72 *(0.119)* *[0.247]*		
β_2 = -1.00	β_2/β_1 = -2.00	-1.41 *(0.24)* *[0.476]*	-1.99 *(0.277)* *[0.277]*	-2.01 *(0.29)* *[0.29]*
β_3 = 2.00	β_3/β_1 = 4.00	2.78 *(0.378)* *[0.868]*	3.92 *(0.365)* *[0.374]*	4.08 *(0.397)* *[0.404]*
γ_0 = -3.50		-2.78 *(0.522)* *[0.89]*		
γ_1 = -1.00		-0.74 *(0.201)* *[0.327]*		
γ_2 = 1.50	γ_2/γ_1 = -1.50	1.02 *(0.435)* *[0.65]*	-1.43 *(0.597)* *[0.601]*	-1.30 *(0.435)* *[0.48]*
γ_3 = 1.00	γ_3/γ_1 = -1.00	0.80 *(0.117)* *[0.233]*	-1.13 *(0.237)* *[0.269]*	-1.07 *(0.226)* *[0.236]*
ρ = 0.15		0.42 *(0.459)* *[0.531]*		

N=2000, 500 simulations. Error terms are jointly standard non-central t-distributed with 4 degrees of freedom and non-centrality parameter 3.
Standard errors in parenthesis and root MSE in brackets.

A second implication is that an increment in sample size reduces the MSE of the semi-parametric ML estimates not only through variance reduction but also through bias reduction. That can be observed in Tables 3.1A and 3.2A for the normally distributed errors, and Tables 3.1B and 3.2B for the non-central t-distributed errors. In both cases the reduction in root MSE is around 30%.

The results of the parametric model under both designs are interesting because they show a high bias in coefficient estimates when normality is erroneously assumed, but a negligible impact on standardized coefficients (ratios to the first slope). Figure 3.1 shows the impact of this result on estimated probabilities and marginal effects for different values of the variable Z_2 in equation 3.8. Z_2 is a continuous variable ranging from 0 to 3, and it is represented in the horizontal axes of figure 3.1. I compare the parametric results with the semi-parametric and the true probabilities using the second design of non-central t-distributed errors. The truth was calculated numerically using a million replications, which explains the reported variability in marginal effects. In general, the consequences of biased coefficients in the parametric model explains differences in estimated probabilities of up to 4 percentage points, while the SMLE produces estimated probabilities much closer to the truth, except for extreme values of Z_2. The same improved performance of SMLE with respect to the parametric model is observed with marginal effects, where the bivariate probit imposes more linearity.

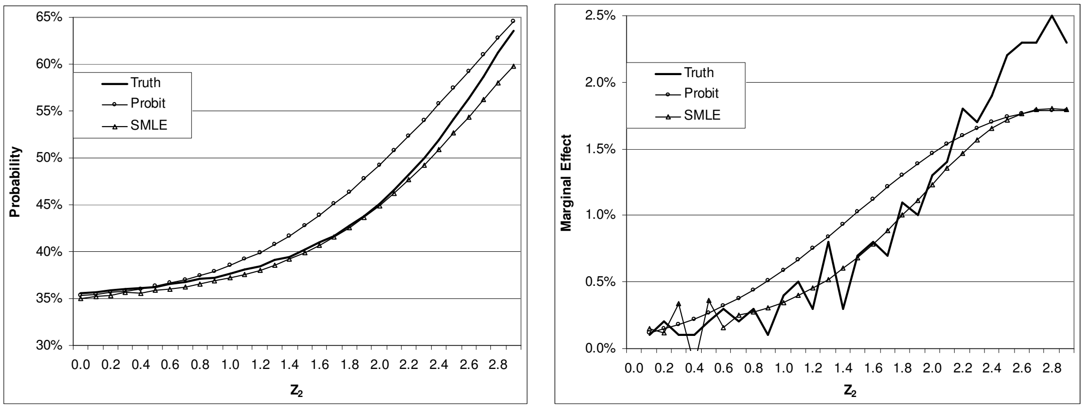

Figure 3.1: Parametric and Semi-parametric Probabilities and Marginal Effects. Non-Normal Errors

For comparison, figure 3.2 reports the same probabilities and marginal effects for the case of normally distributed errors. Clearly, the parametric model produces estimated probabilities and marginal effects that are very close the truth, which is consistent with the results in Table 3.1A. However, the estimated probability from the SMLE is also close to the truth in particular at the mid percentiles of variable Z_2. Bias at extreme values of Z_2 is more problematic for marginal effects.

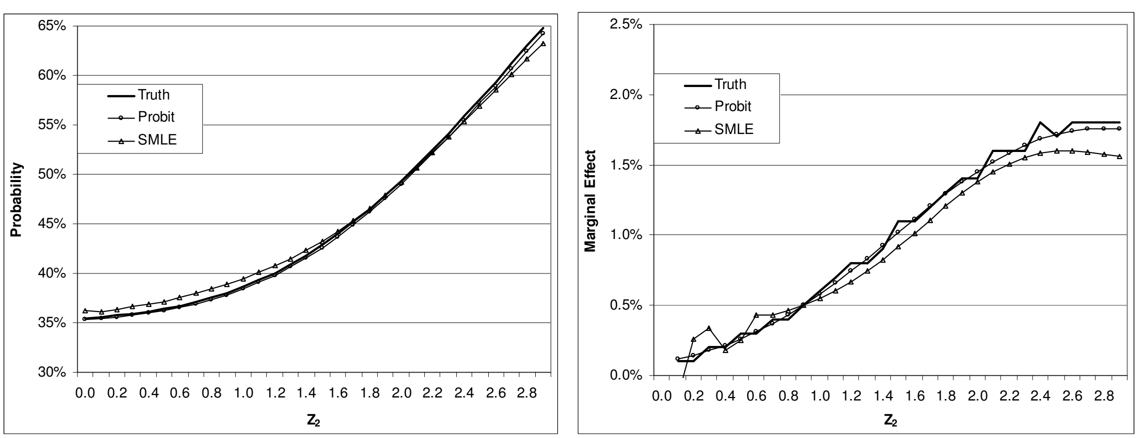

Figure 3.2: Parametric and Semi-parametric Probabilities and Marginal Effects. Normal Errors

The results of this Monte Carlo further justify the use of semi-parametric estimation in the case of the structural misclassification model. The accuracy of the SML estimates in terms of MSE increases with larger sample size through both a reduction in bias and a reduction in variance. The parametric assumption on error distribution may create an important misspecification bias that remains even with larger sample size. Even though it is not the goal of this chapter to test for the parametric assumption, a rough comparison of predicted probabilities, marginal effects and parameter ratios between the parametric and semi-parametric estimation may suggest the magnitude of the misspecification error.

3.4 Comparing Parametric and Semi-parametric Estimation: C-sections in New Jersey

Estimation of double-index SML is highly costly in terms of computer time. Since the cesarean section data in New Jersey accounts for around 100,000 observations every year, in this chapter I use a random sample of 8,777 observations for year 2000, which gives a margin of error of 1%, and a confidence level of 95%. The data correspond to Hospital Patient Discharges collected by the New Jersey Department of Health and Senior Services. These data contain detailed information on each discharge from an acute care hospital including identification of the hospital, patient demographics and zip code of residence, diagnosis and surgical procedures classified by ninth revision of the International Statistical Classification of Diseases and Related Health Problems(ICD-9) codes and Diagnosis Related Group numbers (DRG), source of admission, and identification of payers. Additional socioeconomic information was collected from the US Census 2000, using the patient's zip code as the key variable for matching. Births were identified by DRG codes 370-375. Cesarean sections were identified by DRG 370-371 or ICD-9 code 74xx excluding 7491. The selected sample includes women aged 15 to 49. I excluded deliveries performed in hospitals that in a particular year had less than 100 births, and also excluded patients with invalid zip codes and patients with missing or invalid responses for the variables of analysis.

Estimation of the parametric and semi-parametric models is reported in Table A3.1 in the Appendix. Compared to the parametric results of Chapter 1, parameter estimates are similar, but the significance level was considerable reduced due to the important drop of the sample size. As a consequence, standard errors were reduced by almost one fourth, implying that some coefficients become non-significant as shown in Table A3.1. To compare the parametric and semi-parametric estimates, the ratio of parameters to the first slope was considered for the parametric model. Even though parameter ratio estimates under the parametric and semi-parametric models are different, most of the signs of relationships are the same, with the exception of abruptio placenta in the health status equation, and physician specialty, out-of-pocket and HMO payment in the incentive equation. However, all these variables are not significant in both models. In general, the null is more frequently rejected in the semi-parametric model than in the parametric model, mainly in the incentive equation. The statistically significant results produced by the semi-parametric estimation are consistent with the literature (Keeler et al., 1997; Aron et al., 1998; DiGiuseppe et al., 2001; Rahnama et al., 2006; Pauly, 1980;

and Aron et al. 2000), including the estimation presented in Chapter 1 that uses a larger sample size covering four years of analysis.

Secondly, I compare both models in terms of estimated probabilities. The parametric and semi-parametric estimators produce very similar results. The average probability of c-section, calculated as the average of estimated individual probabilities at actual values, is 23.49% in the parametric model and 23.46% in the semi-parametric model, while the standard error of estimated probabilities is 24.9% in both cases. Figure 3.3 presents a more detailed inspection of probabilities by deciles, sorted on the basis of the estimated semi-parametric probabilities. The figure shows small differences in each decile indicating a reduced misspecification error when bivariate probit is used.

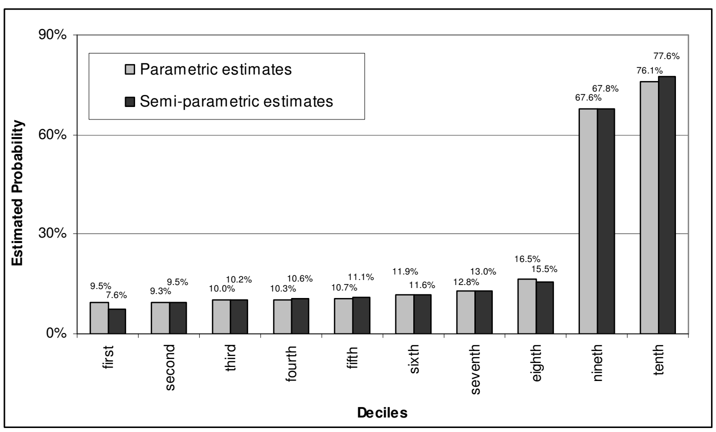

Figure 3.3: Estimated probabilities by deciles: Parametric and semi-parametric estimates

An important result of the misclassification is α_0, the probability of over-treatment or misclassification defined in equation 3.3. In the parametric case, α_0 is estimated using the marginal probability of c-sections due to non-clinical factors after integrating out the clinical factors. Because errors are normally distributed, the marginal probability is also normal, and therefore, it is straightforward to get $\hat{\alpha}_0$. For this sample, the average estimated α_0 is 3.44%[13]. In the semi-parametric case, α_0 is estimated from equation 3.4 after estimating the semi-parametric probabilities $\Pr(y=0)$ and $\Pr(H \geq 0)$[14]. For this method, the average estimated α_0 is 3.90%, which is close to the parametric estimates.

In terms of prediction, the semi-parametric model is only slightly better than the parametric (see Table 3.3). Error type I and II accounts for 14.29% of the data in the parametric case and 14.17% in the semi-parametric case. The difference is negligible and the prediction capability of both models is outstanding.

Table 3.3: Actual and Predicted c-section (in percentages)

Parametric

	Y=1	Y=0
Pred Y=1	14.08	5.00
Pred Y=0	9.29	71.63

Semi-parametric

	Y=1	Y=0
Pred Y=1	14.44	5.24
Pred Y=0	8.93	71.39

Total observations are 8,777. Y=1 is observed c-section, and Pred Y=1 is estimated c-section considering a threshold of 0.5 for predicted probabilities.

[13] This value is very close to α_0 estimated using the full sample of c-sections in New Jersey in period 1999-2002. In Chapter 1 I estimate α_0 equal to 3.2%.

[14] The semi-parametric marginal probability was calculated using the fact that:

$$g(x) = \int f(x, y)dy = \int \nabla_{xy} F(x, y)dy = \nabla_x \int \nabla_y F(x, y)dy = \nabla_x F(x, \infty), \text{ and}$$

$$G(t) = \int_{-\infty}^{t} g(x)dx = \int_{-\infty}^{t} \nabla_x F(x, \infty)dx = F(t, \infty)$$

Finally, I compare both methods in terms of marginal effects. In this case the differences are important. Figure 3.4 shows the estimated probability of c-section and corresponding marginal effect as age, a continuous variable ranging from 15 to 48 years old in the health status equation, increases. The differences in probabilities increase at extreme values of age reaching 1.8% at age 48. At mid values (around 29 years old) the difference is small. The greater linearity in the bivariate probit imposes an important restriction on the parametric marginal effect, explaining the big difference between both methods. The U-shape of the semi-parametric marginal effect is consistent with medical literature that recognizes the greater probability of c-sections in young mothers and elderly mothers. In that regard the semi-parametric estimator captures the non-linearity of age[15].

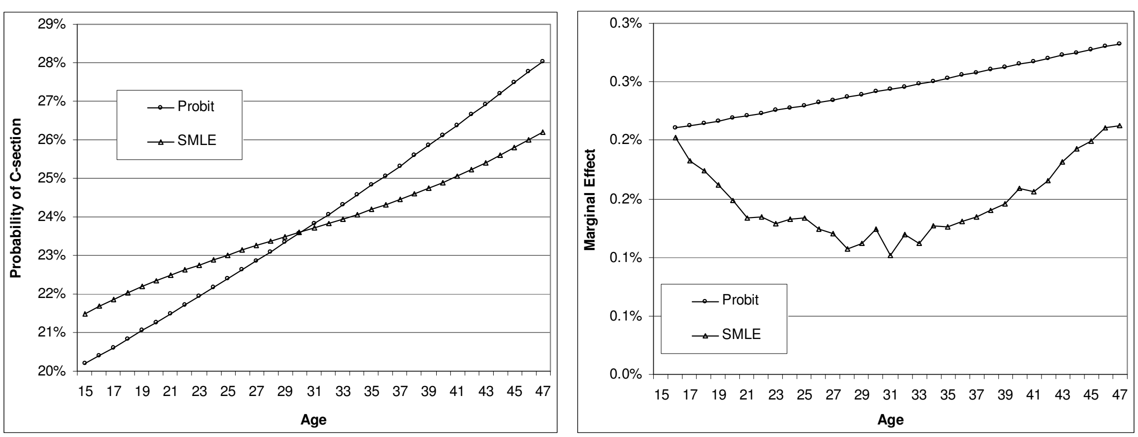

Figure 3.4: Cesarean Section and Maternal Age: Parametric and Semi-parametric Probabilities and Marginal Effects

In figure 3.5 I compare the estimated probability and marginal effect in c-section of a change in employment status of woman, a dummy variable in the incentive equation.

[15] To try to capture the non-linearity of age, the parametric model was re-estimated adding age^2. Under this new parametric specification, age and age^2 were not statistically significant. Further investigation is required to capture the non-linearity of age in the parametric framework.

The figure shows probabilities evaluated at full time employment equal 1 (woman is full time employed) and equal 0. There is a small discrepancy in terms of probabilities and a negligible difference in terms of marginal effects. The change in probabilities due to full employment is 2.75% and 2.19% for the parametric and semi-parametric estimators respectively.

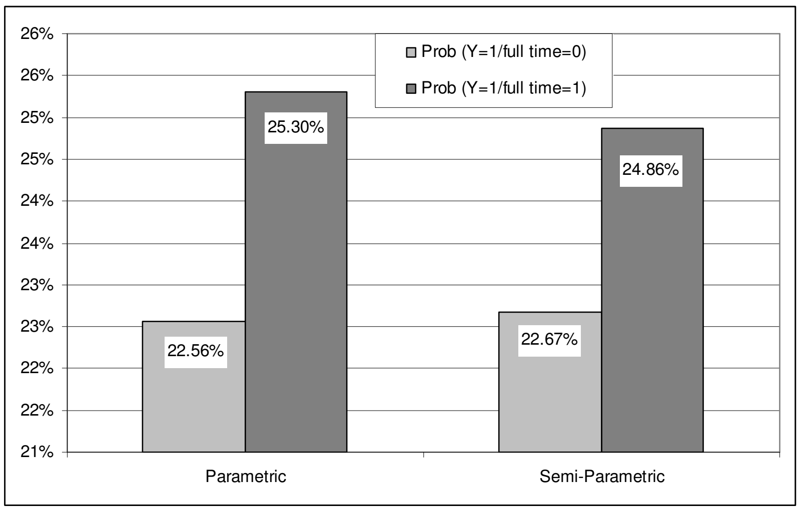

Figure 3.5: Cesarean Section and Woman Employment Status: Parametric and Semi-parametric Probabilities and Marginal Effects

In summary, the parametric and semi-parametric models applied to the c-section delivery in New Jersey show a small difference in terms of probability levels, the probability of over-treatment and prediction. For this particular application, the normality assumption used in the parametric estimation does not appear to create a big inconsistency problem due to misspecification error. Probability estimates are robust to the parametric specification. However, there are important differences in terms of marginal effect (e.g. age) and some parameter ratios. The analysis of these differences requires further investigation.

3.5 Conclusion

This chapter explores the misspecification error in a parametric structural misclassification model of over-treatment. In this special case, the model reduces to a bivariate probit with partial observability. Through a Monte Carlo study I found that misspecification creates an important bias that cannot be reduced even with larger sample size.

Motivated by this result, a semi-parametric estimation is suggested and evaluated. Because of its special characteristics, this model of over-treatment can be estimated by double-index semi-parametric maximum likelihood estimation with partial observability. I used the two-stage bias correction method that Klein and Vella (2008) developed for a double-index SMLE with full observability, which the authors prove to produce root-n consistent and asymptotically normal estimators. The results of the SMLE confirm the consistency of the estimator under different error distributions, and show an important bias reduction when sample size is increased.

In the empirical application, the SMLE is used to estimate over-treatment in cesarean section deliveries, and these results are compared to the parametric estimation to explore the magnitude of misspecification error in the c-section study. The results show a small difference in terms of probability levels and marginal effect for discrete variables and continuous variables at mid values. For this particular application, the misspecification error does not appear to have an important impact on the consistency of parametric estimation. However, there are important differences in terms of marginal effect (e.g. age) and some parameter ratios. These differences together to a full implementation of the semi-parametric method to study the determinants of over-treatment and the only-health-related c-section level in New Jersey using a larger and more representative sample require further investigation.

CONCLUSION

In this dissertation I develop an econometric method to estimate over- and under-utilization of medical procedures. When a physician has incentives that keep him from choosing the appropriate treatment for a patient, the patient's health status loses correspondence with the observed treatment. This generates a problem whose characteristics and effects on estimation are analogous to a classification error. However, this particular measurement error is not random. This chapter proposes a structural model where the classification error is characterized by a physician behavior structure. That allows us to consistently estimate risk-adjusted utilization rates based on clinical factors only, and the probability of inappropriate treatments based on non-clinical factors (misclassification probability). Both measures can be neatly separated to test over- or under- healthcare utilization.

The model is applied to over-treatment of cesarean sections. The first application is to the case of New Jersey in years 1999-2002. The results show that around 3.2% of healthy, non-risky women had c-sections due to non-clinical factors. This rate implies that each year nearly 2,500 women have c-sections for non-medical reasons implying an excess cost of around $17.5 millions per year. Finally, it is estimated that non-clinical factors explain the rapid growth of c-section rates observed in New Jersey over these years.

The second application is to the case of Peru in years 1991-2005. I show that the reform of the Peruvian private health system has increased doctor's financial incentives to overuse medical procedures. After the reform, managed healthcare organizations got enough market power to push healthcare prices down, but were unable to increase the access to private health. With lower incomes and a stagnant market, doctors have higher incentives to increase cesarean sections because they are more profitable than vaginal deliveries. This explains why, after the reform, the c-section rate in the private sector has grown from 27% to 48%, but this rate reaches 66% when women have access to private insurance. In the public sector, however, the c-section rates remain in around 19%. This study shows that each year more than 13 thousand Peruvian women have c-sections without medical justification in the private sector. It represents more than 6.7 million dollars per year paid in excess.

Because the structural misclassification model heavily relies on parametric assumptions about the distribution of unobserved error terms, I explore how robust are estimates of the structural misclassification model are to distributional assumptions. A Monte Carlo study reports that misspecification error might create important bias in the parametric estimation. Motivated by this result, the structural misclassification model is extended to be estimated by semi-parametric methods by using a double-index assumption. A Monte Carlo study of the semi-parametric method confirms the consistency of the estimator under different error distributions, and show an important bias reduction when sample size is increased. I applied the semi-parametric estimation to a sub-sample of cesarean sections in New Jersey. The results show a small difference in terms of probability levels between the parametric and semi-parametric models, which indicate that for this data probabilities are robust to model specification. However, differences in marginal effects and a full implementation of the semi-parametric method to study the determinants of over-treatment and the only-health-related c-section level using a larger and more representative sample require further investigation.

The results of the applied section give direction for further research. A deeper analysis will be done using non-public data related to physician's and hospital's characteristics to understand the main drivers of physician incentives. Additionally, more complete clinical data will be incorporated to measure risk-adjusted utilization rates at the level of hospitals and physicians. A logit version of the model will be developed to be comparable with results from the health service research literature. Finally, a full implementation of semi-parametric estimation will be also considered.

APPENDIX: ESTIMATION RESULTS

Table A1.1: Model estimation of cesarean section deliveries. New Jersey 1999-2002

	Simple Binary Model with Controls (1)	Structural Misclassification Model with independent errors (2)	Structural Misclassification Model (3)
Clinical variables			
Age	0.009 *	0.010 *	0.010 *
	(0.001)	(0.001)	(0.001)
Previous cesarean delivery	1.856 *	1.982 *	1.950 *
	(0.009)	(0.011)	(0.010)
Multiple gestation	0.487 *	0.517 *	0.507 *
	(0.027)	(0.023)	(0.022)
Admission by emergency	-0.207 *	-0.290 *	-0.277 *
	(0.015)	(0.018)	(0.017)
Long labor	0.244 *	0.251 *	0.248 *
	(0.034)	(0.030)	(0.030)
Elderly primigravida (35+ years old)	0.518 *	0.605 *	0.590 *
	(0.031)	(0.026)	(0.025)
Breech or transverse lie presentation	1.702 *	1.830 *	1.799 *
	(0.012)	(0.012)	(0.011)
Diabetes	0.209 *	0.241 *	0.235 *
	(0.015)	(0.014)	(0.014)
Hypertension	0.107 *	0.121 *	0.118 *
	(0.017)	(0.016)	(0.015)
Pre-eclampsia	0.063 *	0.067 *	0.065 *
	(0.025)	(0.024)	(0.023)
Oligohydramnios	0.085 ***	0.055	0.057
	(0.067)	(0.062)	(0.061)
Polyhydramnios	0.692 *	0.764 *	0.748 *
	(0.044)	(0.035)	(0.035)
Abruptio placenta	0.115 *	0.082 **	0.083 **
	(0.042)	(0.039)	(0.038)
Full or partial placenta previa	1.357 *	1.476 *	1.448 *
	(0.060)	(0.046)	(0.046)
Intercept	-1.510 *	-1.669 *	-1.659 *
	(0.023)	(0.023)	(0.023)

Patient and Physician related variables			
Woman is married	-0.042 *	-0.123 *	-0.111 *
	(0.008)	(0.020)	(0.019)
Woman is full time employed	0.153 *	0.419 *	0.417 *
	(0.007)	(0.021)	(0.021)
White (non-Hispanic)	-0.035 *	-0.121 *	-0.115 *
	(0.008)	(0.022)	(0.021)
Black (non-Hispanic)	0.017 ***	0.150 *	0.131 *
	(0.011)	(0.026)	(0.025)
Hispanic	0.063 *	0.129 *	0.131 *
	(0.010)	(0.025)	(0.024)
Zip code mean household income	-0.002 *	-0.004 *	-0.004 *
	(0.000)	(0.000)	(0.000)
Patient payment (uninsured)	-0.139 *	-0.449 *	-0.459 *
	(0.013)	(0.056)	(0.057)
Medicaid payment	-0.068 *	-0.169 *	-0.175 *
	(0.012)	(0.030)	(0.031)
HMO payment	-0.021 *	-0.073 *	-0.070 *
	(0.007)	(0.017)	(0.017)
Yearly average of births in Hospital	0.010 *	0.026 *	0.027 *
	(0.002)	(0.005)	(0.005)
Ob/Gyn Physician	0.031 *	0.157 *	0.163 *
	(0.011)	(0.037)	(0.037)
Year 2000	0.052 *	0.135 *	0.144 *
	(0.009)	(0.027)	(0.027)
Year 2001	0.106 *	0.197 *	0.228 *
	(0.009)	(0.027)	(0.027)
Year 2002	0.177 *	0.366 *	0.397 *
	(0.009)	(0.027)	(0.027)
Intercept	—	-2.058 *	-2.125 *
	—	(0.067)	(0.068)
Correlation	—	—	-0.422 *
	—	—	(0.018)
Degree of physician's incentives	—	0.034	0.032
(Mean of marginal probability)	—	(0.012)	(0.012)
Log-Likelihood function	-159941.87	-160362.97	-160307.28
Number of Observations	403660	403660	403660

Dependent variable is mode of delivery. 1 if it was a cesarean section, 0 if it was a vaginal delivery.

Estimation was done in GAUSS. Program code is available under request.

Standard errors in parenthesis.

* Significant at 1%. ** Significant at 5%. *** Significant at 10%

Table A2.1: Structural Misclassification coefficient estimates

Variables	Before Reform	After Reform
Variables related to monetary and non-monetary incentives		
Birth was in a private facility	0.434 *	0.417 *
	(0.111)	(0.123)
Birth in private sector and private insurance	0.274***	0.267
	(0.196)	(0.273)
Birth was Saturday	0.125	-0.226**
	(0.112)	(0.136)
Birth was Sunday	-0.379**	-0.442**
	(0.208)	(0.203)
Birth was non-working holiday	-0.759**	-0.118
	(0.531)	(0.239)
Variables related to perceived information and preferences of mother		
Woman grew up in a rural area	-0.208**	-0.260 *
	(0.118)	(0.094)
Woman with post-secondary education	0.268 *	0.147**
	(0.093)	(0.086)
Woman decided not to breastfeed	0.538 *	1.053 *
	(0.222)	(0.312)
Household wealth: 1(lowest) to 5(highest)	0.336 *	0.229 *
	(0.072)	(0.051)
Constant	-2.918 *	-1.933 *
	(0.375)	(0.259)
Control variables: Clinical characteristics		
Age (in years)	0.043 *	0.050 *
	(0.005)	(0.015)
Women is 35 years old or more	0.187 *	0.322**
	(0.073)	(0.155)
Multiple birth	0.936 *	1.532 *
	(0.144)	(0.262)
Nulliparity	0.212 *	0.343**
	(0.054)	(0.155)
Number of children (in numbers)	-0.132 *	-0.173 *
	(0.015)	(0.042)
Fever during labor	0.279 *	0.033
	(0.061)	(0.156)
Convulsions during labor	0.160**	0.087
	(0.08)	(0.185)
Under-weight newborn	0.318 *	0.560 *
	(0.064)	(0.142)
Over-weight newborn	0.404 *	0.593 *
	(0.066)	(0.158)
Woman with history of interrupted pregnancy	0.171 *	0.354 *
	(0.046)	(0.106)
Woman had 1 to 3 prenatal visits	-3.382	0.307**
	(9.383)	(0.135)
Constant	-2.182 *	-2.695 *
	(0.142)	(0.523)
Number of observations	9352	4231
Pseudo R-square	0.249	0.284

Dependent variable is c-section birth. Before reform corresponds to period 09/1991-06/1999. After reform corresponds to period 07/1999-09/2005. Standard errors in parenthesis.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table A3.1: Model estimation of cesarean section deliveries. New Jersey 2000

Variables	Parametric Model Estimates	2-Stage BC Semiparametric Model Estimates
	Coefficients	Ratios
Intercept	-1.80*	
	(0.15)	
Age (years)	1.33*	
	(0.41)	
Breech or transverse lie presentation	1.91*	0.78***
	(0.08)	*(0.33)*
Diabetes	0.17**	0.12**
	(0.1)	*(0.07)*
Hypertension	0.26*	0.12***
	(0.11)	*(0.08)*
Pre-eclampsia	0.03	0.06
	(0.17)	*(0.06)*
Oligohydramnios	0.19	0.24*
	(0.41)	*(0.1)*
Polyhydramnios	0.62*	0.27**
	(0.25)	*(0.17)*
Multiple gestation	0.45*	0.35**
	(0.17)	*(0.17)*
Previous cesarean delivery	1.85*	0.70*
	(0.07)	*(0.3)*
Abruptio placenta	0.09	-0.24*
	(0.23)	*(0.06)*
Full or partial placenta previa	0.99*	0.46*
	(0.36)	*(0.2)*
Elderly primigravida >=35 y.o.	0.42**	0.20***
	(0.22)	*(0.14)*
Long labor	0.48*	0.29**
	(0.2)	*(0.16)*
Admission by emergency	-0.64*	-0.18**
	(0.14)	*(0.11)*

Patient and Physician related variables		
Intercept	-1.85*	
	(0.34)	
Zip code mean household income (thousands)	-0.31	
	(0.32)	
Yearly average of births in Hospital (thousands)	0.24	-0.23**
	(0.31)	*(0.13)*
Woman is married	-0.16	0.44*
	(0.13)	*(0.06)*
Obs&Gyn Physician	0.02	0.20*
	(0.19)	*(0.08)*
Woman is full time employed	0.38*	-0.75*
	(0.11)	*(0.06)*
Out-of-pocket payment	-0.23	-0.17
	(0.23)	*(0.14)*
Medicare/aid payment	-0.05	0.03
	(0.17)	*(0.07)*
HMO payment	0.00	0.24*
	(0.11)	*(0.05)*
White (non-Hispanic)	-0.10	0.31*
	(0.16)	*(0.09)*
Black (non-Hispanic)	0.36**	-0.42*
	(0.17)	*(0.05)*
Hispanic	0.28**	-0.12**
	(0.17)	*(0.07)*
Correlation	-0.51***	
	(0.32)	

Dependent variable is mode of delivery. 1 if it was a cesarean section, 0 if it was a vaginal delivery.

Estimation was done in GAUSS. Standard errors in parenthesis.

* Significant at 1%. ** Significant at 5%. *** Significant at 10%

BIBLIOGRAPHY

Abowd, John and H. Farber, 1982. Job Queues and the Union Status of Workers. Industrial and Labor Relations Review 35: 354-367.

Abrevaya, Jason, and J. Hausman, 1999. Semiparametric Estimation with Mismeasured Dependent Variables: An Application to Duration Models for Unemployment Spells. Mimeo.

Alcázar, Lorena and Raúl Andrade. 2000. Transparencia y rendición de cuentas en los hospitales públicos: el caso peruano. Instituto Apoyo, Lima. Documento de Trabajo Nº1.

Alves, Bernadette and A. Sheik. 2005. Investigating the relationship between affluence and elective cesarean sections. BJOG: an International Journal of Obstetrics and Gynecology 112: 994-996.

Amemiya, Takeshi, 1985. Advanced Econometrics. Harvard University Press.

Aron, D., D. Harper, L. Shepardson and G. Rosenthal. 1998. Impact of risk-adjusting cesarean delivery rates when reporting hospital performance, Journal of American Medical Association 279: 1968-1972.

Aron, D., H. Gordon, D. DiGiuseppe, D. Harper and G. Rosenthal. 2000. Variations in risk adjusted cesarean delivery rates according to race and health insurance. Medical Care 38: 35-44.

Arrow, Kenneth, 1963. Uncertainty and the Welfare Economics of Medical Care. American Econonomic Review 53: 941-73.

Belizán, José, F. Althabe, F. Barros and S. Alexander. 1999. Rates and implications of cesarean sections in Latin America: ecological study. British Medical Journal 319: 1397-1402.

Braschi, Roxana. 2005. Las Cesáreas: Un problema del que no se habla. Mimeo.

Brugha, Ruairí and S. Pritze-Aliassime. 2003. Promoting safe motherhood through the private sector in low- and middle-income countries. Bulletin of the World Health Organization 81: 616-623.

Carbajal, Juan Carlos and Pedro Francke. 2000. La Seguridad Social en Salud: Situación y Posibilidades. PUCP, Documento de Trabajo 187.

Chassin, M.R., R.W. Galvin, and the National Roundtable on Health Care Quality. 1998. The Urgent Need to Improve Health Care Quality. Journal of the American Medical Association 280: 1000-5.

Christilaw, J.E. 2006. Cesarean section by choice: Constructing a reproductive rights framework for the debate. International Journal of Gynecology and Obstetrics 94: 262-268.

Cutler, David and Richard Zeckhauser. 2000. The anatomy of health insurance. In Handbook of Health Economics, A.J. Culyer and J.P. Newhouse (Eds.), Volume 1, Part 1, 563-643.

Das, Mitali. 2002. Is there evidence against the induced demand hypothesis? Explaining the large reduction in cesarean rates. Columbia University. Discussion Paper 0102-40.

De Jaegher, Kris and M. Jegers, 2001. The Physician-Patient Relationship as a Game of Strategic Information Transmission. Health Economics 10: 651-668.

DiGiuseppe, D., D. Aron, S. Payne, R. Snow, L. Dierker and G. Rosenthal. 2001. Risk Adjusting Cesarean Delivery Rates: A comparison of hospital profiles based and medical record and birth certificate data. Health Services Research 36: 959-977.

Dranove, David, 1988. Demand Inducement and the Physician/Patient Relationship Economic Inquiry 26: 281-298.

Du Bois, Fritz. 2005. Programas Sociales, Salud y Educación en el Perú: Un Balance de las Políticas Sociales. IPESM, Lima.

Ford, Earl and R. Cooper, 1995. Racial/ethnic differences in health care utilization of cardiovascular procedures: a review of the evidence. Health Service Research 30: 237-252.

Fuchs V, M. McClellan M and J. Skinner, 2001. Area differences in utilization of medical care and mortality among U.S. elderly. NBER Working Paper No. 8628.

Fuchs, Victor R., 1978. The Supply of Surgeons and the Demand for Operations. NBER Working Paper No. 0236.

García N., Luis. 2001. Reforma de la Seguridad Social en Salud en el Perú: Un Análisis Comparativo. PUCP, Documento de Trabajo 196.

Gomes, U., A. Silva, H. Bettiol and M. Barbieri. 1999. Risk Factors for the increasing cesarean section rate in southeast Brazil: A comparison of two births cohorts, 1978 - 1979 and 1994, International Journal of Epidemology 28: 687-694.

González-Rossetti, Alejandra and T. Bossert. 2000. Enhancing the Political Feasibility of Health Reform: A Comparative Analysis of Chile, Colombia, and Mexico, LACHSR, Health Sector Reform Initiative No36.

Gowrisankaran, G. and R. Town 2003. Competition, Payers, and Hospital Quality. Health Services Research 38: 1403-1421.

Gray, R., MA Quigley, C. Hockley, JJ Kurinczuk, M. Goldacre, and P. Brocklehursta. 2007. Cesarean delivery and risk of stillbirth in subsequent pregnancy: a retrospective cohort study in an English population. BJOG An International Journal of Obstetrics and Gynaecology 114: 264-270.

Gregory, Kimberly, L. Korst, M. Krychman, P. Cane, and L. D. Platt. 2001. Variation in Vaginal Breech Delivery Rates by Hospital Type. Obstetrics & Gynecology 97: 385-390.

Gruber, J. and M. Owings. 1996. Physician financial incentives and cesarean section delivery. RAND Journal of economics 27: 99 - 123.

Guzmán, Alfredo. 2002. Para Mejorar la Salud Reproductiva, en La salud peruana en el siglo XXI: Retos y propuestas de política. Juan Arroyo (Ed.). CIES, Diagnóstico y Propuesta 10.

Hanvoravongchai, P., J. Letiendumrong, Y. Teerawattananon, V. Tangcharoensathien. 2000. Implications of Private Practice in Public Hospitals on the Cesarean Section Rate in Thailand. Human Resources Development Journal 4: 2-12.

Hausman, J., J. Abrevaya, and F.M. Scott-Morton, 1998. MisclassiÞcation of the dependent variable in a discrete-response setting. Journal of Econometrics 87: 239-269.

Hueston, W and A. Sutton. 2000. Managed Care Market Share and Cesarean Section rates in united states: is there a link?. American Journal of Managed Care 6: 1202-1208.

Ichimura, Hidehiko and L. Lee, 1991. Semiparametric least squares estimation of multiple index models: Single equation estimation. In Nonparametric and semiparametric methods in econometrics and statistics. Barnett, Powell and Tauchen (Eds.). Cambridge University Press.

Ichimura, Hidehiko, 1993. Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-Index Models. Journal of Econometrics 58: 71-120.

ICPD (International Conference on Population and Development). 1994. ICPD Program of action. Cairo, Egypt.

Iezzoni, Lisa (Ed.), 2003. Risk Adjustment for Measuring Healthcare Outcomes. Third edition. AcademyHealth/HAP.

Keeler, D., R. Park, R. Bell, G. Spelliscy and J. Keesey. 1997. Adjusting Cesarean Delivery Rates for Case mix. HSR: Health Services Research 32: 511-528.

Kenkel, Donald, D. Lillard and A. Mathios, 2004. Accounting for misclassification error in retrospective smoking data. Health Economics 13: 1031-1044.

Kessler, D. and M. McClellan. 2000. Is hospital competition socially wasteful?. The Quarterly Journal of Economics 115: 577-615.

Klein R.W. and C. Shen, 2008. Bias Corrections in Testing and Estimating Semiparametric, Single Index Models. Rutgers University. Mimeo.

Klein R.W. and F. Vella, 2008. Semiparametric Bivariate Selection. Rutgers University. Mimeo.

Klein R.W. and R.H. Spady, 1993. An Efficient Semiparametric Estimator of Binary Response Models. Econometrica 61: 387-421.

Kressin, Nancy and L. Petersen, 2001. Racial Differences in the Use of Invasive Cardiovascular Procedures: Review of the Literature and Prescription for Future Research. Annals of Internal Medicine 135: 352-366.

Leape, Lucian, J. Weissman, E. Schneider, R. Piana, C. Gatsonis, and A. Epstein, 2003. Adherence to Practice Guidelines: The Role of Specialty Society Guidelines. American Heart Journal 145: 19-26.

Lee, Lung-fei, 1995. Semiparametric maximum likelihood estimation of polychotomous and sequential choice models. Journal of Econometrics 65: 381-428.

Leeb, Kira, A. Baibergenova, E. Wen, G. Webster, and J. Zelmer. 2005. Are There Socio-Economic Differences in Cesarean Section Rates in Canada?. Healthcare Policy 1: 48-54.

Lewbel, Arthur, 2000. Identification of the Binary Choice Model with Misclassification. Econometric Theory 16: 603-609.

Li, T., G. Roads, J. Smulian, K. Demissie, D. Wartenberg and L. Kruse, 2003. Physician Cesarean Delivery rates and risk adjusted perinatal outcomes. Obstetrics & Gynecology 101: 1204-1212.

Lin, Herng-Ching and S. Xirasagar. 2004. Institutional Factors in Cesarean Delivery Rates: Policy and Research Implications. The American College of Obstetricians and Gynecologists 103: 128-136.

Luthy, D., J.Malmgren, R.Zingheim, and C.Leininger. 2003. Physician contribution to a cesarean delivery risk model. American Journal of Obstetrics and Gynecology 188: 1579-1587.

MacDorman, Marian, E. Declercq, F. Menacker, and M. Malloy Michael. 2006. Infant and Neonatal Mortality for Primary Cesarean and Vaginal Births to Women with No Indicated Risk. Birth 33: 175-182.

Magder, Laurence and J. Hughes, 1997. Regression when the Outcome is Measured with Uncertainty. American Journal of Epidemiology 146: 195-203.

McGuire, Thomas G. 2000. Physician agency. In Handbook of Health Economics, A.J. Culyer and J.P. Newhouse (Eds.), Volume 1, Part 1, 461-536.

Meng, Chun-Lo and P. Schmidt, 1985. On the Cost of Partial Observability in the Bivariate Probit Model. International Economic Review 26: 71-85.

Meredith, Sheena. 2005. Policing Pregnancy: The Law and Ethics of Obstetric Conflict. Ashgate Pub., USA.

Mesa-Lago, Carmelo. 2005. Las reformas de salud en América Latina y el Caribe: su impacto en los principios de la seguridad social. CEPAL, Documentos de Proyectos No.63.

Minkoff , Howard, K. Powderly, F. Chervenak and L. McCullough. 2004. Ethical Dimensions of Elective Primary Cesarean Delivery. Obstetrics & Gynecology 103: 387-392.

Mossialos, E., S. Allin, K. Karras and K. Davaki. 2005. An investigation of cesarean sections in three Greek hospitals, European Journal of Public Health 15: 288-295.

Murray, Susan. 2000. Relation between private health insurance and high rates of cesarean section in Chile: qualitative and quantitative study. British Medical Journal 321: 1501-1505.

NIH (National Institutes of Health). 2006. State-of-the-Science Conference Statement: Cesarean Delivery on Maternal Request, Obstetrics and Gynecology 107: 1386-1397.

Patel, R., T. Peters, D. Murphy and ALSPAC study team. 2005. Prenatal Risk Factors for Caesarean Section, International Journal of Epidemology 34: 353-367.

Pauly, Mark, 1980. Doctors and Their Workshops: Economic Models of Physician Behaviour. NBER monograph.

Peaceman, A., J. Feinglass and L. Manheim. 2002. Risk adjustment of cesarean delivery rates: a practical method for use in quality improvement. American Journal of Medical Quality 17: 113-117.

Perez-Escamilla, Rafael, I. Maulen-Radovan, and K. Dewey. 1996. The Association between Cesarean Delivery and Breast-Feeding Outcomes among Mexican Women. American Journal of Public Health 86: 832-836.

Poirier, Dale, 1980. Partial Observability in Bivariate Probit Models. Journal of Econometrics 12: 209-217.

Potter, J.E., E. Berquo, I. Perpetuo, O. Leal, K. Hopkins, M. Souza, and M. Formiga. 2001. Unwanted Cesarean Sections among Public and Private Patients in Brazil: Prospective Study. British Medical Journal 323: 1155-1158.

Rahnama, P., S. Ziaei, and S. Faghihzadeh. 2006. Impact of early admission in labor on method of delivery. International Journal of Gynecology and Obstetrics 92: 217-220.

Robinson, R. and A. Steiner. 1998. Models and techniques of managed care, Chris Ham, ed., Managed Health Care: US Evidence and Lessons for the National Health Service, Open University Press, Buckingham.

Salinas, Hugo, S. Carmona, J. Albornoz, P. Veloz, R. Terra, R. Marchant, V. Larrea, R. Guzmány, and L. Martínez. 2004. ¿Se puede reducir el índice de cesárea? Experiencia del hospital clínico de la Universidad de Chile. Revista Chilena Obstet. Ginecol. 69: 8-13.

Schneider, Eric, L. Leape, J. Weissman, R. Piana, C. Gatsonis, and A. Epstein, 2001. Racial Differences in Cardiac Revascularization Rates: Does Overuse Explain Higher Rates among White Patients?. Annals of Internal Medicine 135: 328-337.

SEPS (Superintendencia de Entidades Prestadoras de Salud). 2002. Análisis de rentabilidad de Planes de Salud de las Entidades Prestadoras de Salud. SEPS, Lima.

Sharpe V.A. and A.L. Faden, 1996. Appropriateness in patient care: A new conceptual framework. Milbank Quarterly 74: 115-138.

Sloan N.L, E. Pinto, A. Callec, A. Langerd, B, Winikoff, and G. Fassihiana. 2000. Reduction of the cesarean delivery rate in Ecuador. International Journal of Gynecology and Obstetrics 69: 229-236.

Stewart-Hall, K. 2000. An analysis of risk factors associated with high rates of cesarean births in three selected northeast Tennessee. Thesis, East Tennessee State University.

Tussing, D. and M. Wojtowycz. 1992. The Cesarean decision in New York State, 1986: Economic and Noneconomic Aspects. Medical Care 30: 529-540.

Tussing, D. y M. Wojtowycz. 1993. The effect of physician characteristics on clinical behavior: Cesarean section in New York State. Social Science and Medicine 37: 1251-1260.

Tussing, Dale and M. Wojtowycz. 1994. Health Maintenance Organizations, independent practice associations, and cesarean section rates, Health Serv. Res. 29: 75-93.

Van Ryn, M. and J. Burke, 2000. The effect of patient race and socio-economic status on physicians' perceptions of patients. Social Science and Medicine 50: 813-828.

Villar, José, E. Valladares, D. Wojdyla, N. Zavaleta, G. Carroli, A. Velazco, A. Shah, L. Campodónico, V. Bataglia, and A. Faundes. 2006. Cesarean delivery rates and pregnancy outcomes: the 2005 WHO global survey on maternal and perinatal health in Latin America. The Lancet 367: 1819-1829.

Wennberg, John E., 2002. Unwarranted variations in healthcare delivery: Implications for academic medical centres. British Medical Journal 325: 961-964.

WHO (World Health Organization). 1985. Appropriate technology for birth. The Lancet 2: 436-437.

World Health Organization, 1985. Appropriate technology for birth. The Lancet 2: 436-437.

Xie, Bin, D. Dilts and M. Shor, 2006. The physician-patient relationship: The impact of patient-obtained medical information. Health Economics 15: 813-833.

CURRICULUM VITAE

Alejandro Arrieta

Education

October 2008	PhD, Economics, Rutgers University
March 2002	M.A., Finance, University of the Pacific, Peru
March 1996	B.A., Economics, Pontifical Catholic University of Peru

Teaching Experience

Fall 2007 Spring/Fall 2005 Fall 2004	Instructor (Econometrics, Introduction to Microeconomics and Introduction to Macroeconomics), Rutgers University, NJ
Jan 2006 – Aug 2007	Instructor (Microeconomics, Development and Economic Growth), Univ. of Piura – Univ. of Applied Sciences, Peru

Research and Professional Experience

Feb 2008 - Jun 2008	Research Fellow, Inter-American Development Bank, Washington DC
Aug 2006 - Aug 2007	Associate Researcher, University of Applied Sciences, Peru
Aug 2004 - Aug 2005	Research Assistant, Institute for Health at Rutgers University (IHHCPAR), New Jersey
Aug 2005 - Aug 2007 Apr 1998 – Aug 2002	Senior Economic Analyst, Superintendence of Banking and Insurance, Peru