Estimating the Impact of Physician Incentives on Healthcare Utilization: A Structural Misclassification Error Model

Alejandro Arrieta

(Job Market Paper)
(draft)

October 9, 2007

Abstract

The over-utilization of medical procedures is currently under big debate in the USA. It is well recognized that in the agency relationship between physicians and patients, informational advantages give doctors an incentive to deviate from the right treatment for patients, incurring over- or under- utilization. However, the empirical consequence of this problem has not been adequately considered. In particular, if physicians may alter the truly required treatment, the observed clinical choice will be misclassified, with consequences for unbiasedness and consistency of estimators. In general, empirical literature that does not consider the misclassification error understates the impact of non-clinical factors on healthcare utilization. Econometric literature has focused mainly on the correction for random misclassification, but the measurement error created within the physician-patient relationship is clearly non-random.

This paper proposes a structural misclassification error model in which the physician behavior is modeled to characterize the structure of the mismeasurement error. The model captures the interaction between physician's incentives and patient's health status, and returns unbiased and consistent estimators. It also lets us identify the degree of physician's incentives to deviate from the true treatment choice (misclassification probability), and to compute risk-adjusted utilization rates based uniquely on clinical factors. We apply our model to the cesarean section deliveries in the state of New Jersey for the years 1999-2002. Under this particular setup the model tests demand inducement of c-sections. Our results show a moderate but growing degree of inducement of around 3.2%. We estimate that the rapid growth of the c-section rate in New Jersey over these years is explained mainly by physician inducement.

I. Introduction

Measurement error in the dependent variable is generally innocuous. However, when the dependent variable is limited as in the binary model case, the measurement error (misclassification), leads to biased and inconsistent estimators. 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; Abrevaya and Hausman, 1999; Lewbel, 2000). However, there are special cases where misclassification is a consequence of the action taken by a decision maker who is able to alter the true -unobserved to the econometrician- outcome. In this case, the misclassification error is non-random and the traditional technique to fix the problem will fail unless we came up with a methodology that incorporates the structure of the measurement error into the estimation process.

The ability of a decision maker to alter the true outcome is common in the health economics field. In particular, the appropriate clinical treatment may be affected by doctor's decisions. Pioneered by Arrow's paper on uncertainty in the healthcare market (Arrow, 1963), the informational inequality in the doctor-patient relationship is now well recognized (McGuire, 2000). The informational advantage of physicians with respect to patient's health status creates incentives to overuse or underuse medical procedures according to specific physician's personal goals.

Two examples are illustrative. The first is related to the demand inducement theory, where fee-for-service pricing creates financial incentives for the physician to induce for unnecessary medical procedures (Fuchs, 1978; Dranove, 1988). A well known case is the cesarean section delivery (Gruber and Owings, 1996; Das, 2002; Brown, 1995; Tussing and Wojtowycz, 1993). A healthy woman should have a vaginal delivery based on her health status. This is the appropriate treatment which is observed by the doctor but unobserved by the econometrician. If financial incentives are strong enough to overcome physician ethics, the doctor will induce woman to have a c-section even though it is not clinically called for. In this case, the appropriate choice is affected by the physician's decision, resulting in a misclassified outcome observed by the econometrician.

A second example is related to racial differences in health care access where one of the most studied cases is the lower access of African Americans to cardiovascular procedures (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 physician has a racial bias he may affect the treatment choice by not suggesting the surgery. In that case, the observed outcome will be misclassified resulting in under-utilized cardiovascular procedures for African Americans.

These kind of problems have been vastly studied in the medical and health services research literature to understand what factors drive physician incentives (ref), how HMO's market penetration affects procedure's utilization rates (ref), why national data show important regional variation in the utilization of standard medical procedures (Wennberg, 2002; Fuchs et al., 2001), etc. Epidemiological literature focused on health quality indicators has also recognized the importance of non-clinical factors. Using the same methodology to compute case-mix risk adjusted mortality rates (ref. iezzoni), this literature has calculated risk adjusted utilization rates controlled for non-clinical factors to report hospital level quality ratings.

However, this literature has not adequately recognized and corrected for the misclassification problem arose by the nature of the doctor-patient relationship. The biasedness and inconsistency generated by the misclassification error has two important impacts: First, the magnitude of the physician's influence on utilization rates has been understated. Second, the statistical significance of some health risk factors has been overstated.

The main contribution of this paper is to use a structural model of physician behavior to characterize and correct for misclassification error. That allows us three improvements compared to previous literature: First, to estimate unbiased and consistent parameters for patient's health risk factors and physician's incentives. Second, to model physician's behavior and its interaction with patient's health status. Third, to be able to estimate the degree of inducement (misclassification probability) and risk adjusted utilization rates based only on health characteristics (removing the inducement). The second section of this document describes the structural misclassification error model and establishes a parametric solution. In the third section we use a Monte Carlo study to compare the effect on estimators' unbiasedness of four different approaches to estimate risk adjusted utilization rates, ranging from no considering the misclassification problem to consider it adequately. The fourth section provides an application to demand inducement in cesarean section deliveries. The last section concludes.

II. A parametric estimation of structural misclassification

Getting back to our two examples, the physician observes the true health condition of his patient. Conditioned on patient's health status, the doctor may deviate from appropriate treatment if his personal goals overcome his professional ethics within his utility function. In that regard, the structure of this model is a simplified version of game theory models of inducement (De Jaeguer and Jegers, 2001; Xie et al. 2006.) The physician's decision tree is shown in Figure 1. In the first stage, nature plays and determines patient's state (h: healthy or sickly), which is only observed by the physician. Two treatments are considered: A and B. We assume that doctor's (monetary or not-monetary) benefits under treatment A are higher than under 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). He may decides to do the right or appropriate treatment for the patient ($i<0$) or the wrong or inappropriate one (i>0) based on patient's health state (h)[1]. One treatment, however, gives more utility to the doctor (for instance, treatment A: c-section for a pregnant woman, or no-cardiovascular surgery for an African American). Therefore, aside from professional ethics, the physician will choose treatment A rather than the right treatment for each patient's health state.

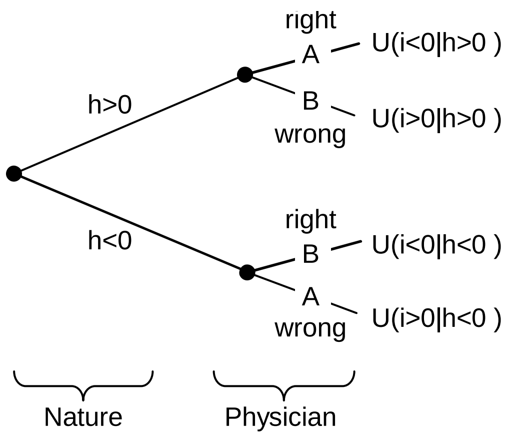


Figure 1: Physician's decision tree

The preferred treatment may be the right treatment depending on patient's health state. Following our two examples, the right treatment for an at-risk pregnant woman will be a c-section that is also the preferred treatment for the physician. The right treatment for a healthy African American will be no-cardiovascular surgery that is also preferred for the physician. Clearly, under this patient's health state, the physician doesn't have any incentive to affect appropriate treatment choice. The implication of this result is that under one patient's health status there is no misclassification error (the right choice is not affected by the doctor). But under the health status where the physician's preferred treatment is the wrong treatment for the patient, the misclassification error shows up. Because the misclassification error works only in one direction, this model is identifiable in terms of the monotonicity condition described by Haussman et al. (1998)[2]. It is important to remark that this model may be easily extended to the case of

[1] Even though it is not explicitly modeled, the interaction between the physician and patient is captured by the index i, that can be interpreted as a net-cost of inducement. The cost is given by doctor's professional ethic and the degree of patient's information and resistance to physician's influence.
[2] The monotonicity condition is needed to identify the parameters, and it requires that the probabilities of misreport sum less than 1. In our case this is automatically satisfied since there is only one probability.

misclassification in both directions, when under- and over- utilization are present in a given treatment. This is, for example, the case of cardiac revascularization, where some authors suggest that instead of under-use for African Americans there is an over-use among white patients (Schneider et al., 2001).

The econometric model is described following figure 1. In the first stage, patient's health status (h) is determined by a set of clinical characteristics or risk factors (x).The physician can observe patient's health condition with error:

$$h = x\beta + \varepsilon_h \qquad (1)$$

This is the health status equation. The patient will require a specific treatment if health status exceeds zero. Econometrician will only observe the treatment choice (y). Without physician incentives to alter the required treatment, the econometrician should observe the choice y=1 if $h \geq 0$, and in that case a binary model estimation will be unbiased and consistent, since the probability of observing and not-observing the treatment choice is respectively

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

However, if physician decides to do the surgery when it is not needed -as in the cesarean section case-, or not to do the procedure when it is required -as in the cardiovascular surgery-, then the binary model estimation will be biased and inconsistent since $\Pr(y = 1) \neq \Pr(h \geq 0)$. 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 (z) which are observed by the doctor with error.

$$i = z\gamma + \varepsilon_i \tag{2}$$

This is the physician's incentives equation. When incentives exceed a threshold 0, the doctor proceeds with the wrong treatment altering the appropriate choice with probability (misclassification probability)

$$\Pr(i \geq 0 \mid h) \tag{3}$$

Therefore, the econometrician observes the treatment choice with a probability that depends on (3). In particular, the probability of observing a c-section and a vaginal delivery is respectively:

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

While the probability of observing a patient with or without cardiovascular surgery is respectively:

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

The parameters β, γ in equations (1) and (2) can be estimated with MLE, maximizing the likelihood function

$$L(\beta, \gamma) = \prod_{i=1}^{n} \Pr(y = 1)^{y} \Pr(y = 0)^{1-y} \tag{4}$$

Notice that if the errors $\varepsilon_h, \varepsilon_i$ are not independently distributed, the problem becomes a bivariate one. Additionally, if it is assumed that the error terms are normally distributed, the problem becomes a sequential probit.[3] In the particular case of the c-section delivery, the likelihood function is

$$L(\beta,\gamma)=\prod_{i=1}^{n}[1-\Phi(-x\beta,-z\gamma)]^{y}[\Phi(-x\beta,-z\gamma)]^{1-y} \qquad (5)$$

Notice that the structure of this model is very similar to the partial observability model (Poirier, 1980). Even though both models end up with the same likelihood function, 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) and (2) may be based on two separate set of variables. A variable may be in both equations if it carries information about both: patient's health status and physician's incentives (some examples are age, sex, weight, etc. depending on the analyzed treatment). That implies that the identification problems that are problematic in the partial observability model are easily satisfied in the structural misclassification model. In particular, identification is easily satisfied when at least two different continuous variables are included in each equation. However, because observability is limited, this model has the same cost than the partial observability model in terms of efficiency of the maximum likelihood estimator (Poirier, 1980; Meng and Schmidt, 1985).

The structural misclassification error model presented in this paper rests on strong parametrical assumptions. We have based our estimation on a bivariate probit, but the model can be easily extended to a bivariate logit. However, a natural extension to rid of the parametric assumptions is to estimate this model semi-parametrically based on a multiple index model as in Ichimura and Lee (1991).

[3] See Amemiya (1985) for a description of those models.

III. Monte Carlo simulation

In this section we discuss the impact of different approaches to estimate risk adjusted utilization rates for medical procedures. The literature has not adequately recognized the potential misclassification error involved in procedures that are over or under used. Even thought it is accepted that non-clinical factors affect health outcomes, literature has either do not consider these variables at all (ref.) or consider them as control variables (ref.) The agency problem in the physician-patient relationship creates important interactions between patient's health status and non-medical factors. Consequently, any approach that does not consider non-medical factors creates an omitted variable bias in addition to the misclassification bias. By controlling by non-medical factors, the bias is reduced but it is still important because the interaction between patient's health status and physician incentives is not taken in account, and the misclassification error is not corrected.

The estimator based on the maximization of the likelihood function (4) considers both, interaction and misclassification correction. A restricted and simplified version of the structural misclassification model is to assume that error terms in equations (1) and (2) are independent. From an empirical perspective this restriction may be strong, since there are many health characteristics that are non-observable to the econometrician but to the physician. Therefore, ε_h may convey information that could affect non-observable physician's incentives (ε_i).

In order to asses the impact of this four approaches on estimators' bias and consistency, I examine the results of Monte Carlo simulation. The design considers 1000 draws of a sample size of 5000. The true model representing equations (1) and (2) is

$$h = -1.5 + 0.5x_1 - x_2 + 2x_3 + \varepsilon_h$$

$$i = -1 - 1.5z_1 + z_2 + 0.5z_3 + 2z_4 + \varepsilon_i$$

Covariates x and z include dummy variables and continues variables drawn from uniform and trimmed chi-squared distributions. The error disturbances $\varepsilon_h, \varepsilon_i$, are drawn jointly from a bivariate standard normal distribution with correlation $\rho = 0.25$. Results of the Monte Carlo study are reported in Table 1.

[Insert Table 1]

The results of the Monte Carlo simulation are consistent with the misclassification problem described by Hausman et al. (1998). In particular, probit understate estimated coefficients. The probit model with only patient's health related variables (model I) produces estimates that are biased by 65-75%. When physician's and patient's characteristics are used as control variables within the traditional probit (model II), bias is slightly reduced in the case of risk health factors. However, coefficient estimates of control variables (non-clinical characteristics) have a lower downward bias of around 15-35%.

The size of the bias in the health status and the non-clinical coefficients depends, among others, on two parameters: the error correlation and the degree of inducement (misclassification). Different Monte Carlo designs (not shown) were used to see the impact of both parameters on estimator's biasedness. First, the bias increases in both set of estimators when the correlation gets closer to 1. Second, the bias in the health status estimators decreases and in the inducement estimators increases when the degree of inducement falls.

Estimators of the structural misclassification model are unbiased and present larger standard errors. Compared with these models, the traditional probit overstates the precision of their estimates. As an implication for epidemiological studies the impact of health risk factors on utilization rates will appear less important than they really are when probit/logit is used. Confidence intervals will be narrower too.

There is a small discrepancy between MLE estimators when error independence is erroneously imposed (model III). In general, estimated coefficients of the restricted MLE are lower and biased by 1-3% in the case of risk factors and physician's and patient's characteristics. For the unrestricted MLE estimation (model IV), the estimators are unbiased (less than 0-1%) and show slightly higher standard errors that the restricted MLE.

An important feature of the structural misclassification model is that it allows us to neatly separate the estimated physician's incentives probability and the utilization rate due to health status only. Table 1 reports the percentage of inducement or degree of misclassification. It was calculated as the marginal probability that physician's incentives exceed the threshold of inducement: $\Pr(i > 0) = \int \Pr(i > 0, h) dh$. Clearly, the estimated probability of inducement is unbiased. The same methodology may be used to get an estimated utilization rate based on health status only, under the counterfactual that there are not physician's incentives to influence the true health outcome.

IV. An application to demand inducement in cesarean section deliveries

This section applies the method to births in the state of New Jersey in the period 1999-2002. We use Hospital Patient Discharge Data collected by the New Jersey's 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's demographics and zip code of residence, diagnosis, surgical procedures, source of admission, and identification of payers. Births were identified by DRG codes 370-375, and a list of procedures were identified using ICD-9 codes. Cesarean sections were identified by DRG 370-371 or ICD-9 code 74xx excluding 7491. Additional socioeconomic information was collected from the US Census 2000, using the patient's zip code as key variable for matching.

Two sets of variables are used in the estimation. The first set is related to woman's health characteristics that have been identified by previous medical research (Keeler et al., 1997;

Aron et al., 1998; DiGiuseppe et al., 2001; and Rahnama et al., 2006). The second set of variables is related to patient and physician characteristics that may drive doctor's incentives. The complete list of variables and their means values for vaginal and c-section deliveries are reported in Table 2.

[Insert Table 2]

In general, variables in the inducement equation may be classified in three types with respect to the impact of patient's information on physician inducement. The first type of variables indicates some degree of resistance to physician's inducement. Here we include two variables that capture the physician's perception about patient-obtained medical information: ethnicity (Black, Hispanic or White woman) and average household income in zip code of patient's residence. It is expected for the doctor that minorities and lower income women have less access to medical information making them more vulnerable to inducement (Xie et al. 2006; Pauly, 1980). Social support measured as the presence of a partner (married or life partner) may indicate a better degree of information, so it is expected that it reduces inducement and therefore the probability of c-section.

The second set of variables may reflect compliance with physician's inducement. Here we consider woman's employment status. It is expected that full employed women may prefer a c-section delivery because of the convenience in terms of scheduling and the lower pre-partum work.

The third set of variables is not related to the degree of patient-obtained medical information. Here we include the woman's insurance condition (uninsured, Medicaid[4], HMO, other private insurance). The type of insurance is important for inducement because it sets the method of payment for the physician. It is expected that compared to non-HMO private insurance, the uninsured have the lowest rate of c-sections since they must afford the payment. The capitation payment of HMO reduces physician incentives,

[4] Few patients were also identified as Medicare beneficiaries, but were included in the Medicaid groups because of the method of payment.

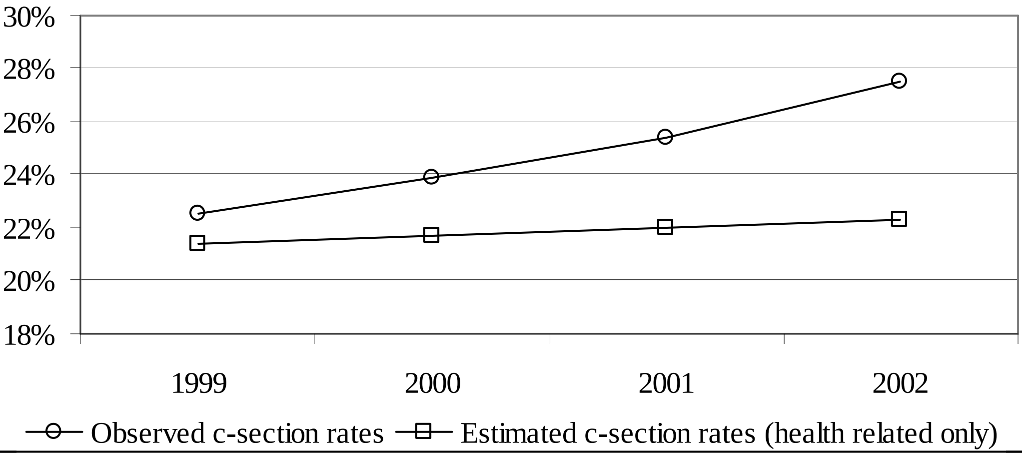
and so thus the prospective payment system under Medicaid. The size of the hospital has been shown to have an impact of over-utilization due to the supply-sensitive service phenomenon (Wennberg, 2002). Here we measure hospital's size as the average yearly number of births observed in each hospital. Physician specialty (Ob/Gyn specialty) captures the tendency of over-utilization in more specialized doctors. Finally, year dummy variables are also included in the estimation.

In this particular application, the uses of patient level data allow us to test for demand inducement. In contrast to estimation with aggregate data, patient level data controls for endogenous changes in the supply permitting a cleaner analyses of demand inducement (McGuire, 2000; Pauly, 1980). In that regard, the approach is similar to Gruber and Owings (1996) who used individual level data to show that the fertility decline in the 1970's reduced ob/gyn income and caused a small but significant increment in cesarean section deliveries. However, Gruber and Owing followed the traditional approach where interactions between woman's health status and misclassification error are not considered. Additionally, as was described in section 3, our approach allows us to measure the level of inducement and the expected c-section rate if inducement were eliminated.

Table 3 reports the estimation using three approaches: The traditional approach (model I), the structural misclassification approach with independent errors (model II), and the structural misclassification approach with dependent errors (model III). Compared to the structural misclassification model with dependent errors, the traditional approach understates both the impact of woman's health characteristics and physician inducement related variables. However, the bias is greater for inducement related variables, and that is explained by the small degree of inducement (misclassification) observed in the data. The difference between model II and model III is small in spite of the high and significant negative correlation in the error terms.

For the whole period, the estimated marginal probability of inducement was 3.2%. It means that 3.2% of non at-risk women had non-clinical required cesarean section due mainly to physician's inducement. It means 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 inducement 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 may be explained by inducement rather than changes in health conditions in the population. To test that hypothesis we compute the marginal probability of c-sections due to health conditions only. This estimated c-section rate is the rate without inducement, and therefore without misclassification. Figure 2 compares both, the observed and the only-health-related estimated c-section rate. According to this figure, c-section rates in New Jersey growth from 22.5% in 1999 to 27.5% in 2002. However, this growth was explained by non-clinical required c-sections mainly. Without physician incentives, the c-section rate in New Jersey would have kept almost constant at around 22%.

Figure 2: Cesarean Section Rates: New Jersey 1999-2002



What are the determinants of inducement? The estimated inducement equation is shown in Table 3. We structure the discussion of these results according to the three types of variables described before, and using marginal calculations for each estimated parameter. We found the expected direction for all variables related to resistance to inducement. Income and ethnicity are observed by physicians as an indicator of patient observed medical information. For the average household income, an increment of 5 thousand dollars reduces the probability of c-section in 0.10%. This small but significant impact was also observed by Pauly (1980) for ambulatory care. Ethnicity have an important impact on probability of c-sections. Black and Hispanic women have respectively 2.20% and 2.30% higher probability of having a c-section related to other no-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 inducement, implying a 1.90% lower probability of c-section.

Compliance to physician's inducement was capture with women employment. As we expected, full employed women are more induced for c-section, increasing the probability of c-section in 7%. Li et al. (2003) show also higher c-section rates for employed women.

The third type of variables is not related to patient's degree of information. The most important and studied variable is the source of payment. With respect to non-HMO private insured patients, uninsured women are the less induced, with reduction in probability of inducement of around 9.20%. Medicaid beneficiaries are the next in lower inducement, with reduction in the probability of c-section of 3.40%. Finally, the capitated payment system of HMO protects against physician inducement, reducing probability of c-section in 1.30%. With respect to hospital size, our 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 raise the probability of c-section in 0.10%. A similar argument is validated when we observe that women attended by more specialized physicians (Ob/Gym) have higher probabilities of c-section (2.60% more).

V. Conclusions

[will be added soon]

VI. References

[will be completed soon]

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

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

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

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.
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.

Ichimura, Hidehiko and Lung-Fei 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.

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

Lewbel, Arthur, 2000. Identification of the Binary Choice Model with Misclassification. Econometric Theory16, 603-609.

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

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

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

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

Schneider, Eric, Lucian Leape, Joel Weissman, Robert Piana, Constantine Gatsonis, and Arnold Epstein, 2001. Racial Differences in Cardiac Revascularization Rates: Does "Overuse" Explain Higher Rates among White Patients?. Annals of Internal Medicine 135, 328-337.

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.

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

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

VII. Appendix

Table 1. Monte Carlo simulations.

	True	Probit model (I)	Probit model with controls (II)	MLE assuming independence (III)	MLE (IV)
Health condition: Risk factors					
β_0	-1.50	0.142	-0.672	-1.578	-1.505
		(0.043)	(0.073)	(0.256)	(0.234)
β_1	0.50	0.171	0.200	0.483	0.495
		(0.133)	(0.146)	(0.304)	(0.309)
β_2	-1.00	-0.274	-0.328	-0.976	-1.010
		(0.042)	(0.046)	(0.202)	(0.195)
β_3	2.00	0.682	0.808	1.948	2.004
		(0.067)	(0.073)	(0.275)	(0.257)
Doctor's and Patient's characteristics					
γ_0	-1.00	–	–	-1.019	-1.014
				(0.159)	(0.142)
γ_1	-1.50	–	-1.031	-1.460	-1.506
			(0.147)	(0.218)	(0.223)
γ_2	1.00	–	0.652	0.970	1.008
			(0.045)	(0.105)	(0.106)
γ_3	0.50	–	0.352	0.490	0.504
			(0.071)	(0.100)	(0.102)
γ_4	2.00	–	1.638	1.974	2.007
			(0.071)	(0.113)	(0.111)
ρ	0.25	–	–	–	0.238
					(0.201)
% of inducement †	0.554	–	–	0.541	0.552
				(0.021)	(0.020)

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

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

Table 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
Health related variables			
Age (years)	28.62	30.47	29.08
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
Multiple gestation	0.71	2.84	1.24
Previous cesarean delivery	4.01	42.00	13.44
Abruptio placenta	0.47	0.60	0.50
Full or partial placenta previa	0.10	0.87	0.29
Elderly primigravida >=35 y.o.	0.69	1.39	0.86
Long labor	0.76	0.78	0.77
Admission by emergency	5.58	3.30	5.02
Patient and Physician related variables			
Woman is married	65.90	71.48	67.28
Zip code mean household income (thousands, $)	56.00	57.15	56.29
Yearly average of births in Hospital (thousands)	2.48	2.61	2.51
Obs&Gyn Physician	89.67	91.09	90.02
Woman is full time employed	34.85	40.46	36.24
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
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
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

Table 3: Model estimation of cesarean section deliveries. New Jersey 1999-2002

	I	II	III
Health related variables			
Age	0.009 *	0.010 *	0.010 *
	(0.001)	(0.001)	(0.001)
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)
Multiple gestation	0.487 *	0.517 *	0.507 *
	(0.027)	(0.023)	(0.022)
Previous cesarean delivery	1.856 *	1.982 *	1.950 *
	(0.009)	(0.011)	(0.010)
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)
Elderly primigravida (35+ years old)	0.518 *	0.605 *	0.590 *
	(0.031)	(0.026)	(0.025)
Long labor	0.244 *	0.251 *	0.248 *
	(0.034)	(0.030)	(0.030)
Admission by emergency	-0.207 *	-0.290 *	-0.277 *
	(0.015)	(0.018)	(0.017)
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)
Zip code mean household income	-0.002 *	-0.004 *	-0.004 *
	(0.000)	(0.000)	(0.000)
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)

Woman is full time employed	0.153 *	0.419 *	0.417 *
	(0.007)	(0.021)	(0.021)
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)
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)
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)
Inducement (Marginal probability)	—	0.034	0.032
Log-Likelihood function	-159941.87	-160362.97	-160307.28
Pseudo-R2	0.523	0.522	0.522
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.

Model I is the traditional probit. Model II is the structural misclassification model assuming independent errors. Model III is the structural misclassification model with dependent errors.

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

Standard errors in parenthesis.

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