Testing exogeneity in the bivariate probit model: Monte Carlo evidence and an application to health economics

Daniele Fabbri, Chiara Monfardini*, Rosalba Radice
Department of Economics, University of Bologna, Italy.

July 16, 2004

Abstract

Many economic applications involve the modeling of a binary variable as simultaneously determined with one of its dycotomous regressors. In this paper we deal with a prominent health economics case study, that of cesarean section delivery utilization across public and private hospitals. Estimating the probability of cesarean section in a univariate framework neglecting the potential endogeneity of the hospital type dummy might lead to invalid inference. Since little is known about the exact sampling properties of alternative statistics for testing exogeneity of a dycotomous regressor in probit models, we conduct an extensive Monte Carlo experiment. Equipped with the simulation results we apply a comprehensive battery of tests to an Italian sample of women and find clear evidence against exogeneity of the hospital type dummy. We speculate on the economic implications of these results and discuss the misleading interpretation arising from the adoption of either univariate probit model or seemingly unrelated bivariate probit model.

Keywords Bivariate probit model, endogenous dummy, exogeneity testing, cesarean section delivery, hospital choice

J.E.L. C15,C35, C52, I11

Acknowledgments
We are grateful to Renzo Orsi and Joao Santos Silva for helpful comments. We also thank the participants to the Computational Econometrics and Statistics Workshop (Neuchatel) and CEMAPRE seminar (ISEG, Lisbon) for stimulating discussions. The usual disclaimer applies.

*corresponding author: Department of Economics, University of Bologna, Piazza Scaravilli, 2 - 40126 BOLOGNA. Italy. e-mail: chiaram@economia.unibo.it.

1 Introduction

Dycotomous indicators, like mortality or treatment choice, are the core outcomes examined in the applied health economics literature. Determinants for the probability of such outcomes typically include -via appropriate dummy variables- qualitative information about the health care provider. In most cases this latter is determined by an individual choice process in which perceived quality and reputation play a major role. This brings about a self-selection of patients into hospitals based on observables and unobservables characteristics that also determine the dycotomous indicator of interest. As a consequence, hospital dummy variables are potentially endogenous in this equation, even after controlling for a full set of patient's covariates. In a recent paper Geweke, Gowrisankaran and Town (2003) propose a Bayesian approach to the modeling of mortality in presence of nonrandom selection of patients to hospitals. Their contribution is aimed to provide an appropriate inference tool for assessing hospital quality with patient discharge databases, a major issue in health policy. On a sample of patients admitted to 114 californian hospitals, they find a strong evidence that the more frail patients patronize the higher quality hospitals, and point out that ignoring nonrandom admission would lead to invalid inference on quality.

We examine here the related issue of "practice variation" across hospital types, i.e. the variation in utilization of procedures for which questions have been raised about overuse, underuse, or misuse.[1] This constitutes an important policy issue for healthcare cost containment and represents the only available alternative for quality assessment whenever mortality, or other adverse medical outcome, is a rare event. Modeling practice variation across hospitals potentially brings a further source of nonrandom selection due to patients unobserved preferences for a given treatment.

In the application of this paper we focus on a very common surgical procedure: cesarean section (CS) delivery. For this procedure the issue of cost containment -due to its overuse- and that of misuse are relevant enough to attract a massive attention in the applied health econometric literature. A typical regressor in binary response models for the probability of deliverying via CS is a dummy conveying information about the ownership structure of the hospital as a proxy for unobservable differences among hospitals. In our case study, referring to Italy, women can choose between two main hospital types: public or private. CS delivery variation observed across these two types is of some interest for the policy maker when dealing with issues of authorization, licensing, and reimbursement. However, in this framework, the problem of nonrandom selection appears as a major concerns, as far as women are allocated into the two hospital types according to unobservable characteristics (related to either clinical conditions or preferences) that might be among the determinants of the

[1] Abnormal utilization rates of vaginal birth after cesarean, laparoscopic cholecystecomy, CABG, PTCA, or bilateral cardiac catheterization are all likely candidates for inappropriate or inefficient delivery of care. See, for instance, the Agency for Healthcare Research and Quality (2002) AHRQ Quality Indicators.

probability of CS. It should be mentioned here that women may differ in their degree risk aversion (either for themselves or for the fetus) and in their aversion to pain and suffering.

This paper contributes to the existing literature on both the health economics and the econometric methodology grounds. We adopt here a bivariate probit model in which the hospital type dummy is allowed to be endogenously determined in the CS probability equation.[2] This represents an improvement on the existing probability models for CS delivery, relying on single-equation estimation as, for example, in Gruber and Owings (1996). In the resulting econometric framework, represented by the so called recursive bivariate probit model (Heckman, 1978, Maddala, 1983), the hypothesis of exogeneity of the dummy can be defined as the absence of correlation between the error terms and submitted to statistical test. For this purpose, alternative approaches can be invoked: Lagrange Multiplier, Conditional Moment Tests, Likelihood Ratio and Wald test. The first two involve only single equation probit estimation, and are particularly attractive for their simplicity, while the second two are based on simultaneous estimation. Application of these inference tools is still rare in the microeconometric practice, although the recursive bivariate probit model is increasingly adopted in different fields[3].

The properties of these tests are only known asymptotically, and there is no evidence on their finite sample performance. Nevertheless, it is well known that inference in binary models is generally demanding in terms of sample sizes. Kooreman (1994) and Smith and Moffat (1999) show that estimation of the correlation coefficient in bivariate binary models implies a considerable loss of information with respect to the fully-observed dependent variable case, so that large samples are needed to get precise estimates. In this paper we investigate the exact sampling properties of a number of exogeneity test statistics through an extensive set of Monte Carlo experiments. As a by-product, we analyse the finite sample properties of the estimator of the endogenous dummy coefficient and the existence of possible "compensating effects" between the latter and the correlation coefficient.

We then apply the whole battery of tests to our case study. We find that different test statistics lead to different conclusions and are able to use the simulation results as a guideline to interpret this finding. We find strong evidence against the hypothesis that the hospital-type dummy is exogenous in the equation determining CS probability. The negative sign of the correlation coefficient suggests that a self-selection mechanism allocating the more risky patients to

[2]In Fabbri and Monfardini (2001) this model is adopted within the context of a test for demand induction on a dataset referred to a Local Health Authority coming from patient discharge database.

[3]In health economics the model has been used to analyse the effect of supplemental insurance ownership on dycotomous health demand indicators (see Holly et al., 1998, Buchmueller, 2002) and to explore endogeneity of self-reported disability measure for the decision to apply for social benefits (Benitez-Silva et al., 2003). In the law and economics field, Deadman and MacDonald (2004) investigate the relationship bewteen criminal behaviour and victimization. In labour economics, Pudney and Shields (2000) and Bryson, Cappellari and Lucifora (2003) adopt the germane framework represented by the recursive ordered probit.

public hospital is prevailing, while, after controling for observable and unobservable women characteristics, the private hospital type dummy has a positive impact on the CS probability. Finally, we discuss the extent to which neglecting the endogeneity issue and limiting the analysis either to the univariate probit framework or to the Seemingly Unrelated Regression Equation (SURE) probit would lead to a misleading interpretation of the results.

The rest of the paper has the following structure. Section 2 introduces the notation for the model, the associated maximum likelihood inference, and the exogeneity test statistics whose sampling behaviour is explored in our simulation study. Section 3 describes the Monte Carlo design and discusses the results of the experiments. Section 4 illustrates our case study, presents the estimates of the model and the outcome of the exogeneity tests. Section 5 contains some final remarks.

2 The bivariate probit model with endogenous dummy

2.1 The model

The bivariate probit model with endogenous dummy model belong to the general class of simultaneous equation models with both continuous and discrete endogenous variables introduced by Heckman (1978). In his systematic review of multivariate qualitative models Maddala (1983) lists this model among the recursive models for dycotomous choice (Model 5). The recursive structure builds on a first reduced form equation for the potentially endogenous dummy - and a second structural form equation determining the outcome of interest:

$$\begin{array}{l} y_{1i}^{*} = \beta_1' x_{1i} + u_{1i} \\ y_{2i}^{*} = \beta_2' x_{2i} + u_{2i} = \delta_1 y_{1i} + \delta_2^{'} z_{2i} + u_{2i} \end{array} , \tag{1}$$

where: y_{1i}^{*} and y_{2i}^{*} are latent variables, y_{1i} and y_{2i} are dycotomous variables observed following the rule:

$$\left\{ \begin{array}{lll} y_{li} = 1 & if & y_{li}^{*} > 0 \\ y_{li} = 0 & if & y_{li}^{*} \leq 0 \end{array} \right. ; l = 1, 2;$$

x_{1i} and z_{2i} are vectors of exogenous variables, β_1 and δ_2 are parameter vectors, δ_1 is a scalar parameter, and $\beta_2' = \begin{array}{cc} (\delta_1 & \delta_2')' \end{array}$ .The error terms are assumed to be independently and identically distributed as bivariate normal:

$$\left(\begin{array}{c} u_{1i} \\ u_{2i} \end{array} \right) \sim IIDN \left(\left[\begin{array}{c} 0 \\ 0 \end{array} \right], \left[\begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array} \right] \right).$$

A widespread opinion in the literature is that the parameters of the second equation are not identified unless x_{1i} includes at least one variable not contained in z_{2i} (as in linear simultaneous equations for fully observed endogenous variables). This assertion, stated by Maddala (1983) is contraddicted in a recent

paper by Wilde (2000), who shows that exclusion restrictions are not needed provided there is one varying exogenous regressor in each equation. Therefore, x_{1i} and z_{2i} are not necessarily distinct regressor vectors. The parameter vector $\underset{Kx1}{\beta} = (\beta_1' \quad \beta_2' \quad \rho)'$ can be estimated by maximum likelihood. The sample loglikelihood function, that has to be maximized resorting to numerical methods[4] is given by:

$$l(\beta) = \sum_{i=1}^{N} \left[d_{11} \ln P_i^{11} + d_{10} \ln P_i^{10} + d_{01} \ln P_i^{01} + d_{00} \ln P_i^{00} \right] \tag{2}$$

where:

$d_{11} = y_{1i}y_{2i}, = y_{1i}(1 - y_{2i}), d_{01} = (1 - y_{1i})y_{2i}, d_{00} = (1 - y_{1i})(1 - y_{2i})$

$P_i^{11} = prob(y_{1i} = 1, y_{2i} = 1 | x_{1i}, z_{2i}) = \Phi_{i2}(\beta_1' x_{1i}, \delta_1 + \delta_2^{'} z_{2i}, \rho)$

$P_i^{10} = \Phi_{i2}(\beta_1' x_{1i}, -\delta_1 - \delta_2^{'} z_{2i}, -\rho),$

$P_i^{01} = \Phi_{i2}(-\beta_1' x_{1i}, \delta_2^{'} z_{2i}, -\rho),$

$P_i^{00} = \Phi_{i2}(-\beta_1' x_{1i}, -\delta_2^{'} z_{2i}, \rho)$

and $\Phi_{i2}(.,.,\rho)$ is the bivariate normal distribution function of the model error terms.

2.2 The exogeneity hypothesis and MLE inference

In the above setting, the exogeneity condition is stated in terms of the correlation coefficient ρ, which can be interpreted as the correlation between the unobservable explanatory variables of the two equations. When $\rho = 0$, y_{1i} and u_{2i} are uncorrelated and y_{1i} is exogenous for the second equation of model 1. On the contrary, $\rho \neq 0$ implies that y_{1i} is correlated with u_{2i} and therefore endogenous. This lead us to the definition of the following hypothesis system, whose testing is the object of our interest:

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

Hereafter we'll refer to the null hypothesis H_0 as to the exogeneity hypothesis. It must be remarked that under H_0 the loglikelihood 2 becomes the sum of the loglikelihood functions of two univariate probit, say:

$$l_0(\beta) = l_1(\beta_1) + l_2(\beta_2)$$

so that the MLE can be written as $\widetilde{\beta}_0 = (\widetilde{\beta}_1' \quad \widetilde{\beta}_2' \quad 0)'$, where $\widetilde{\beta}_1$and $\widetilde{\beta}_2$ are obtained by separate ML estimation of the corresponding univariate probit equations. On the other side, under H_1, simultaneous estimation is needed to get consistent estimates of the parameters of the second equation $\beta_2' = (\delta_1 \quad \delta_2')'$ .The great simplification of the model occurring under H_0 makes reliable testing tools

[4] As first and second order derivatives have "simple" analytical expressions (see Greene, 2000) Newton's iterative algorithm can be adopted to solve the maximization problem.

for such hypothesis particularly appealing to the applied econometrician. The aim of next sections is to review different test statistics suitable for this scope. and to provide evidence on their absolute and relative finite sample performance.

2.3 Alternative approaches to exogeneity testing

A number of procedures is available for testing exogeneity in simultaneous equation models involving limited dependent variables, including the two-stage tests proposed by Smith and Blundell (1986) for tobit models with continuous endogenous explanatory variables and by Rivers and Vuong (1988) in the probit context. A thourough review is given in Maddala (1995). Here, we focus on the presentation of the four main testing approaches that are suitable in probit models when the potentially endogenous variable is dycotomous, as they are still rarely - but increasingly considered by the practiotioners, and no evidence exist about their sampling properties. These can be grouped in two classes: Lagrange multiplier (LM) and conditional moment (CM) tests are particularly attractive as they are based on univariate probit estimation of the two equations, while likelihood ratio test and the "t" test based on MLE of ρ (which gives the Wald test when squared) rely on simultaneous estimation of the bivariate probit. We sketch hereafter the main computational features of the four testing procedures,

2.3.1 LM Tests

Define the score vector and the information matrix of our model as:

$$d(\beta) = \frac{\partial l(\beta)}{\partial \beta} = \begin{bmatrix} \frac{\partial l(\beta)}{\partial \beta_1} \\ \frac{\partial l(\beta)}{\partial \beta_2} \\ \frac{\partial l(\beta)}{\partial \rho} \end{bmatrix}$$

$$I(\beta) = -E\left[\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta'}\right] = E\left[d(\beta)d'(\beta)\right]$$

The LM test statistic for $H_0 : \rho = 0$ is given by:

$$LM = d(\widetilde{\beta}_0)' I(\widetilde{\beta}_0)^{-1} d(\widetilde{\beta}_0) \underset{H_0}{\overset{d}{\rightarrow}} \chi_1$$

Evaluation in $\widetilde{\beta}_0$ simplifies the score vector to: $d(\widetilde{\beta}_0)' = \begin{bmatrix} 0 & 0 & \frac{\partial l(\beta)}{\partial \rho} \end{bmatrix}_{\beta=\widetilde{\beta}_0}$

It is well known that different estimation methods for $I(\beta)$ lead to different versions of the LM test. A first choice lies in the evaluation of the analytical expetcation, a second one relates to the use of the information equality. The so-called OPG variant involves only first order derivatives. Its simplest form, avoiding analytical expectation, is given by:

$$I_{OPG}(\beta) = G^{'}(\beta) G(\beta),$$

where the $i-th$ row of matrix $G(.)$ evaluated in $\widetilde{\beta}_0$has the simplified expression:

$$G_i(\widetilde{\beta}_0) = \left[\begin{array}{ccc} \frac{\partial l_{1i}(\beta_1)}{\partial \beta_1'} & \frac{\partial l_{2i}(\beta_2)}{\partial \beta_2'} & \frac{\partial l_i(\beta)}{\partial \rho} \end{array} \right]_{\beta=\widetilde{\beta}_0}$$

Hereafter we list alternative implementations of the LM test we consider in this study, dividing them into two main groups according to their computational features.

1. The first version of the LM test is obtained through the residual sum of squares (RSS) in the artificial regression (Davidson and MacKinnon, 1984, 1989):

   $$i = G(\widetilde{\beta}_0)\gamma + \eta$$

   and it is evaluated as $LM1 = N - RSS$. This is a computational device to obtain the quantity:

   $$LM1 = \left[\frac{\partial l(\widetilde{\beta}_0)}{\partial \rho}\right]^2 V_{LM1}$$

   where V_{LM1} is the (K, K) element of $I_{LM1}^{-1}(\widetilde{\beta}_0) = \left[G'(\widetilde{\beta}_0)G(\widetilde{\beta}_0)\right]^{-1}$. Therefore, LM1 uses OPG variant for estimating the information matrix, without taking the analytical expectation.[5]

2. Kiefer (1982) investigates the form of the LM statistic for testing dependence in multivariate SURE probit models. He shows that, under H_0, $p\lim \frac{\partial^2 l}{\partial \rho \partial \beta} = 0$, so that asymptotically the information matrix is block diagonal. We denote the LM test statistic relying on this simplification as:

   $$LM^{Kifier} = \left[\frac{\partial l(\widetilde{\beta}_0)}{\partial \rho}\right]^2 V^{Kiefer}$$

   with $V^{Kiefer} = -E\left[\frac{\partial^2 l}{\partial \rho \partial \rho}\right]^{-1}$. We consider three different evaluation methods of V^{Kiefer}, leading to three further versions of the LM statistic. All these implementations of the LM test adopt a diagonal matrix,to estimate the information matrix, imposing a simplification which is only valid asymptotically and they differ from LM1 which is based on the inversion of a "full" information matrix.

   - $LM2$ uses OPG form for estimating V^{Kiefer} and does not take anayt-ical expectation
   - $LM3$ uses OPG form and evaluates analytical expectation (this is the version suggested by Greene's textbook (2000) for bivariate SURE probit)

[5] In principle the results of Davidson and MacKinnon to compute hessian-based versions of LM misspecification tests statistics in univariate probit models through appropriate artificial regressions could be extended to the test of the exogeneity hypothesis.

- $LM4$ uses second order derivatives and does not take anaytical expectation

2.3.2 CM Test

This general approach to misspecification testing has been proposed by Newey (1985) and Tauchen (1985), and surveyed by Pagan and Vella (1989) with focus on individual data models. Skeels and Vella (1999) investigate the performance of CM tests of various forms of misspecification in univariate probit and tobit models and evidence a genearl poor performance of alternative version of the tests in the probit context. The CM test for the hypothesis of exogeneity is based on the statement that under H_0 the moment condition $E_0[u_{i1}u_{i2}] = \underline{0}$ has to hold. Its sample counterparts can be evaluated using the pseudo-residuals obtained by MLE estimation of the bivariate probit under H_0, $\widetilde{\beta}_0 = (\ \widetilde{\beta}_1' \quad \widetilde{\beta}_2' \quad 0\ )'$:

$$\widehat{\tau} = \frac{1}{N}\sum_{i=1}^{N} \widehat{u}_{1i}\widehat{u}_{2i}$$

with $\widehat{u}_{li} = \widehat{E}\left(u_{li}|y_{li}\right) = \widehat{\phi}_i \left(y_i - \widehat{\Phi}_i\right) \widehat{\Phi}_i^{-1} \left(1 - \widehat{\Phi}_i\right)^{-1}$, $l = 1, 2$. The CM statistic we consider in our study is equivalent to the (squared) "t- test statistic" for $\alpha_2 = 0$, asymptotically distibuted as $N(0, 1)$, in the artificial regression:

$$i = C(\widetilde{\beta}_0)\alpha_1 + \widehat{\tau}\alpha_2 + \nu$$

where:

$$C_i(\widetilde{\beta}_0) = \left[\ \frac{\partial l_{1i}(\beta_1))}{\partial \beta_1'} \quad \frac{\partial l_{2i}(\beta_2))}{\partial \beta_2'}\ \right]_{\beta_1 = \widetilde{\beta}_1, \beta_2 = \widetilde{\beta}_2}$$

This version corresponds to choosing OPG form to evaluate the information matrix entering in the expression of the asymptotic covariance matrix of the test.

2.3.3 Likelihood ratio and Wald tests

The further test statistics we include in our investigation are those requiring estimation of the model under the alternative hypothesis. The likelihood ratio test has the well known form:

$$LR = -2\left[l(\widetilde{\beta}_0) - l(\widehat{\beta})\right] \underset{H_0}{\overset{d}{\rightarrow}} \chi_1$$

while the "t-test" based on ρ, whose square gives the Wald test, is given by:

$$RHO = \frac{\widehat{\rho}}{se(\widehat{\rho})} \underset{H_0}{\overset{d}{\rightarrow}} N(0, 1)$$

require estimation of $se(\widehat{\rho})$. We use to this purpose the corresponding element of the inverse of the negative hessian matrix $V = -E\left[\frac{\partial^2 l_i(\beta)}{\partial\beta\partial\beta'}\right]$.

3 Monte Carlo analysis

3.1 Design of the experiments

Our sampling experiments are based on the bivariate probit model:

$$\begin{cases} y_{1i}^* = \beta_{10} + \beta_{11} x_i + \beta_{12} z_i + u_{1i} \\ y_{2i}^* = \delta_{10} y_{1i} + \delta_{11} y_{1i} z_i + \delta_{20} + \delta_{21} z_i + u_{2i} \end{cases} \tag{3}$$

where we insert in the second equation of the "classical" model 1 an interaction term between the endogenous dummy and the exogenous regressor z. This generalization is meant to reflect the frequent case in which the dycotomous variable has an impact not only on the intercept of the latent equation, but also on its slope. It is evident that if $\delta_{11} \neq 0$, but the multiplicative dummy term is omitted, the additive term becomes correlated with the resulting error component and this might affect the outcome of the exogeneity test. This is one of the items we examine in the subsequent Monte Carlo analysis.

We set (x_i, z_i) as standard normal variables with correlation 0.5, while the error terms (u_{1i}, u_{2i}) are generated, according to the assumption of the model. We consider three different Data Generating Processes (DGP) characterized by the following parameter sets:

	β_{10}	β_{11}	β_{12}	δ_{10}	δ_{11}	δ_{20}	δ_{21}
DGP1	0.5	1	1.5	1	1	-0.5	0.5
DGP2	-1.5	0.5	0.5	1.5	-1	-0.5	-1
DGP3	-1.75	0.7	0.4	-0.7	0.6	1.9	-1

The resulting sampling schemes originates pairs of observable binary variables with different marginal probability distributions across the two outcomes. In DGP1 both binary variables have "balanced" distribution on the two outcomes, while DGP2 and 3 represents more realistic cases in which the potentially endogenous variable is less frequently observed with value 1.[6] Therefore DGP1 is the easiest to estimate, while DGP2 and 3 will require more information. As a consequence, we set the sample size $N = 500, 1000$ in DGP1, while DGP2 and 3 are examined with sample sizes $N = 1000, 2000$.Empirical size and power results of the test are obtained by varying the correlation coefficient as follows: $\rho = 0, \pm 0.25, \pm 0.50, \pm 0.75$.

[6] The two-way frequency tables, evaluated for samples of 100000 observations, with $\rho = 0.5$ exhibit the following pattern in the three DGPs:

	DGP1			DGP2			DGP3		
	y_2=0	y_2=1		y_2=0	y_2=1		y_2=0	y_2=1	
y_1=0	0.36	0.60	0.42	0.54	0.33	0.87	0.08	0.82	0.90
y_1=1	0.08	0.50	0.58	0.08	0.05	0.13	0.04	0.06	0.10
	0.44	0.56	1	0.62	0.38	1	0.12	0.88	1

3.2 Monte Carlo results

We discuss in this section the most salient features of the simulation experiment results we obtained with 5000 replications. The corresponding detailed tables - 1 to 5 - are contained in Appendix 1. The results have been obtained using the simulation setup of STATA 8.2. To obtain MLE of the bivariate probit model, we resorted to the command "biprobit", which exploits the Newton-Raphson maximization method and allows for Hessian-based estimation of the asymptotic covariance matrix. Such command, presented in STATA for the SURE bivariate probit, sorts out the correct estimation procedure also when one of the dependent dycotomous variable is included as a regressor for the other probit equation, as the two models share the same log-likelihood "mechanics".[7]

3.2.1 Empirical size

Table 1.1 in Appendix 1 provides rejection frequencies generated under $H_0 : \rho = 0$ of the seven test statistics outlined in the previous section based on the 10%, 5%, 1% critical values of the corresponding asymptotic distributions. These size results share some important common features across the three DGPs. Starting from the statistics evaluated through simultaneous estimation, we note that LR sistematically outperforms the other tests for all values of N and different nominal levels. On the other side RHO tends to displays high over-rejection patterns, especially in DGP2 and DGP3 (the most difficult to estimate), where the over-rejection is still serious with $N = 2000$. We will discuss later the findings of our investigation about the source of the bad performance of the RHO test in small samples. Turning to the tests based on single equation estimation, two features are worth being remarked. First, CM1 and LM1 - which give results very similar to each other- display empirical sizes quite close to the nominal ones, performing only slightly worse than LR. This is an important result for the practitioners, as it indicates that they can count on reliable inference tool for testing exogeneity without coping with simultaneous estimation. This finding is encouraging and particularly relevant in the context of multivariate probit models more complicated than the bivariate, where avoiding joint estimation is even more desirable. Second, the set of LM tests (LM2, LM3, LM4), originated by the Kiefer formulation of the test statistic variance, displays very unsatisfactory finite sample sizes, with zero rejection frequencies for the analyzed sample dimensions. Therefore these formulas should be applied with enormous caution. We recall that while LM1 uses a full matrix to estimate the information matrix, the other LM tests adopt a diagonal matrix (whose inversion leads to different results from the full one), imposing a simplification which is only valid asymptotically. We find here that this simplification is crucial with respect to other choices available for computing the information matrix, like analytical expectation evaluation or OPG form. A final comment concerns the effect on the

[7] This has been verified through comparison with an ad hoc routine. The STATA "do files" implementing the exogeneity test statistics presented in the previous section are available from the authors upon request.

size of the tests due to erroneously omitting the multiplicative dummy term in model 3. In DGP1 and DGP2 the introduced size distortion turned up to be negligible, while the figures reported in Table 1.2 show that in DGP3 such a specification error makes empirical sizes of all the tests (except for LM2, LM3, LM4) almost double than in the corresponding correctly specified model case reported in the bottom panel of Table 1.1.

3.2.2 Empirical power

For all DGPs we evaluated two set of results. The first set, denoted by nominal power, refers to rejection frequencies obtained under a set of alternative hypotheses ($\rho = \pm 0.50, \pm 0.75$) using theoretical critical values. For brevity sake, we do not report these results as the observed different power performances are influenced by the different size patterns of the various test statistics (with RHO displaying highest rejection frequencies of the true alternative hypothesis, LM2,LM3,LM4 the lowest). Tables 2.1 to 2.3 contains instead exact powers, i.e. rejection frequencies based on the finite sample distribution critical values obtained from the previous size experiments.[8] We observe that after the critical values corrections the set of test statistics considered give similar and satisfactory results in all thre DGPs, including the LM2, LM3, LM4 statistics. A further pattern that can be noticed inside each DGP is the asimmetry of the powers for values of ρ differing in sign with the same absolute values. For a given sample size, DGP1 and DGP2 display higher power against alternative hypotheses characterized by positive values of ρ, while in DGP3 the pattern is reversed. Take for example the LR test at 5% level. When $N = 1000$, the exact power is about 94% in DGP1, 36% in DGP2, 42% in DGP3 for $\rho = 0.5$, while for $\rho = -0.5$, we obtain 93% in DGP1, 27% in DGP2, 55% in DGP3. When $N = 2000$ the exact power is about 99% in DGP2, 98% in DGP3 for $\rho = 0.75$, while for$\rho = -0.75$, we get 80% in DGP2, 100% in DGP3. From these figures it is also evident that differences can be spotted across the three sampling schemes for given values of N and ρ, with DGP1 displaying highest power, followed by DGP3 and DGP2. This ordering reflects the increasing amount of sample size information required for the inference in the three considered generating schemes.

3.2.3 Finite sample performance of the endogenous dummy and the correlation coefficients

In this section we exploit the simulation experiments output to evaluate the finite sample properties of MLE of the two crucial parameters of the model, i.e. the potentially endogenous dummy coefficient δ_2 and the correlation coefficient ρ. It might be argued in fact that it is empirically difficult to disentangle the direct effect of y_1 on y_2 from the effect exerted on the latter variable through the correlation of the error terms u_1 and u_2. In Table 3.1 to 3.2 we report some indicators of the finite sample distributions of $\widehat{\rho}$ and $\widehat{\delta}_2$. Looking at the correlation coefficient results, it can be noticed that the general performance of

[8] Results on nominal power are available on request.

MLE (for both the coefficient and the standard error of its estimate) strongly depends on the features of the DGP. In DGP1 the estimators display means across the replications that are close to the true values, with moderate standard deviations that appears to be well estimated through the asymptotic covariance matrix formula (cf. the second and third column of the Tables). In DGP2 and DGP3, we observe substantial deviations of sample means from corresponding true values, and higher standard deviation of the estimates. Moreover, the asymptotic covariance matrix formula delivers us estimated standard errors that understate the empirical ones. When $\rho = 0$ this underestimation pattern can explain the pronounced over-sized behaviour displayed by the RHO test statistic for DGP2 and DGP3.

Turning to the dummy variable coefficient, it can be noticed that in DGP1 and DGP2 δ_2 is estimated with satisfactory precision, at the sample dimensions considered, for different values of ρ, while in DGP3 δ_2 is estimated with low precision, also when $N = 2000$. A similar comment can be expressed looking at Table 4, where we report the empirical second type error probabilities of the null hypothesis $H_0 : \delta_2 = 0$, i.e. the probability of erroneously concluding that the dummy coefficient is equal to zero when it is not in the various DGPs. Comparing the results across different values of ρ with those obtained when $\rho = 0$ reveals that such probability might be high (especially for DGP3) at the considered sample sizes N, but this behaviour is independent from the fact that we are estimating the correlation coefficient. This provides evidence that, according to the structure of the process generating the data, a lot information might be required for estimating with precision the dummy coefficient, regardless the estimation of ρ. This does not rule out the possibility that a sort of compensating effect might exist between the two estimates $\widehat{\rho}$ and $\widehat{\delta}_2$, making it difficult to perform reliable finite sample inference on the two coefficients. We explore this issue in the following section.

3.2.4 Misspecification of the discrete choice model

This final part of our simulation experiment, presented in Table 5, looks at the consequences of misspecifying the discrete choice model on the inference on δ_2 and ρ. To get some flavour of this impact, we generated three datasets using the DGP1, DGP2, DGP3 schemes, setting $\rho = 0.5$ and $N = 5000$. We then estimated on the three artificial datasets three alternative model. The first estimated model is a univariate probit for the second equation, which erroneously imposes $\rho = 0$. A serious distortion of $\widehat{\delta}_2$ emerges for the three DGPs. In particular, such a coefficient is over-estimated, taking up the effect of the correlation coefficient. The second specification is a seemingly unrelated bivariate probit model for both equations, which eliminates the structure by erroneosly imposing $\delta_2 = 0$. This results in distorting the correlation coefficient estimate, with a direction which depends on the true sign of δ_2. Finally, the last column of Table 5 contains estimated parameters of the correctly specified bivariate probit with endogenous dummy. The model appears to be able to disentangle the two coefficients, that are well estimated in DGP1 and DGP2, while DGP3

is more problematic, although the null hypotheses corresponding to the true values could not be rejected. This piece of evidence confirms the existence of the above mentioned compensating effects between the two estimates, which turns out not to be a major problem in our particular setting. Nevertheless, we recommend that in the empirical analysis, a lot of caution is invoked whenever it is observed that the dummy coefficient is significantly different from zero in univariate probit estimation, while bivariate probit estimation leads to unsignificant estimates of both the dummy and the correlation coefficient. This pattern can interpreted as a signal of the existence of an identification problem for the parameters of the model, indicating that more information is required.[9]

4 An application to probability of Cesarean Section delivery with endogenous hospital type dummy

As we mentioned in the introduction we deal here with the analysis of CS delivery utilization across two classes of hospital, public and private, on a nationally representative sample of Italian women in childbirth. In the Italian National Health Service the two types of hospitals differ in several respects. Public hospitals are run by Local Health Authorities (LHA) or by autonomous public trusts and are mainly financed through fixed budgets. Privately licensed hospitals can treat patients within the NHS, i.e. free of charge, and are afterwards refunded by the LHA which the patient belongs to. Private hospitals' refunding is based on the prospective payment of each clinical episode (classified into DRGs). Public and private hospitals differ also in terms of quality and infra-structural capacity. For our case study in delivery we notice, for instance, that public hospitals do have emergency surgical capacity and newborn intensive care units.[10] Private do not have emergency room and therefore are not allowed to admit on an emergency. Finally, looking at the style of practice, the presence of teaching personnel could reasonably increase the role of professional and deontic rewards in the public leading to a higher propensity to improve clinical practices and to adopt the more appropriate ones. Patients are completely free to choose the treating hospital; it may be public or privately licensed, both within or outside their assisting LHA or region. Since patients are totally unaware of treatment costs, choice is essentially determined by distance from home, hospital reputation, waiting lists, and perceived quality.

Given these differences our main aim is to establish which part of the observed variation in CS utilization (see table 1) across the two mentioned types of hospital is due to a pure hospital type effect and which one to patients self-sorting into them.

[9] We experimented this kind of identification problem in past trials with other datasets, and observed it in some of the existing applications.

[10] WHO (1985) recommends that "natural deliveries after a caesarean should normally be encouraged wherever emergency surgical capacity is available"

Table 1: Cesarean section incidence. (standard deviations) in parentheses.

	CS incidence	**"Market Shares"**
Public hospital	28.4 (.451)	91.8
Private hospital	38.9 (.488)	8.2
All	29.2 (.455)	100.0

The econometric framework we adopt is that described in model 1 with the following adaptations:

$y_{1i} = 1$: if woman chooses a private hospital to deliver,

$y_{2i} = 1$: if woman delivers via CS.

x_{1i}, z_{2i} contain socioeconomic and relevant risk factors for CS delivery.

The two variables are simultaneous determined as far as we can not exclude that common unobservables enter both equations or that the unobservables of one equation are correlated with those of the other. As we mentioned in the introduction, two "opposing" self-selection mechanisms might be in place: one related to tastes (women with idiosyncratic preferences for CS delivery are more likely to choose an hospital more willing to comply with their preferences), the other related to clinical conditions (women with unobserved indications for a difficult delivery are more likely to go to an "high quality" hospital). Incidentally we aim at suggesting which of these two conflicting mechanisms is prevailing in our sample.

4.1 Data description

We work on a dataset coming from the "Indagine Statistica Multiscopo sulle Famiglie: condizioni di salute e ricorso ai servizi sanitari" (ISMF), a national household survey conducted by the Italian National Institute of Statistics (ISTAT) every 5 years. The last available survey was conducted from september 1999 to august 2000 when a sample of 40119 households were interviewed. The survey provides a full account of individual health condition, health care utilization, biometric parameters plus socio-economic status (education, working condition) and other relevant economic variables like complementary private health insurance holdings. In this study we exploit a section of the survey focussing on the last delivery experienced by female components of each sampled household in the five years before the survey. Delivery experience is described in an individual self-compiled part of the survey. Therefore data about mode of delivery, health problems suffered and theraphies underwent during pregnancy and delivery are self-reported. We have 5660 women filling in this section of the survey for a corresponding number of deliveries. However, for data coeherency, we decided to use only those delivering in the two years before the survey. We therefore ended up with a sample of 3498 delivering women.

From the ISMF we were able to obtain a full set of control variables (see table 2) comprising individual predisposing risk factors for CS delivery, some

socioeconomic variables, and residential area controls. Given the self-compiled nature of the questionnaire our set of risk factors do not include most of the clinical conditions usually controlled for in the health econometrics analysis of CS variation (see for example Dubay, Kaestner, and Waidmann, 1999). Major lacks are controls for breech presentation, prior CS, and fetal distress.

4.2 Main results

Table 3 presents the main findings emerging from the following specifications: univariate probit, seemingly unrelated bivariate probit, and bivariate endogenous dummy model. A full account of our estimation exercise is available in Appendix 2.[11] For the sake of brevity we only notice here that overall results are coherent with expected signs. Each risk factor contributes to increase the probability of CS delivery; while they are almost uniformly not significant in driving hospital choice. Socioeconomic variables (education and insurance holding in our final specification, but also family income in a previous estimation) seem to be irrelevant in determining CS probability. However, the woman being self-employed, holding a private health insurance and being more educated has a higher probability to deliver in a private hospital. According to this evidence neither of the two self-selection mechanisms we referred above operate through our observables. This does not preclude that the same does occur through unobservable variables.

In the univariate setting we notice that, after controlling for confounders, women going to a private hospital significantly receive more CS than their counterfactuals going to a public hospital. For a representative women with no risk factor and average levels for age and socioeconomics this amounts to a .069 increase in the CS probability. We might be tempted to ascribe such an increase to systematic differences in the style of practice and in the incentives structure across the two types.[12] Being less ambitious we can anyway draw some relevant policy implications. Given that the overuse of CS delivery is an indication of bad quality in hospital healthcare provision, we could conclude that, according to the univariate model, private hospitals are of lower quality than public ones. However this implication strongly relies upon the hypothesis of exogeneity of the hospital dummy. This actually motivated our careful analysis of the issue. We therefore applied the complete battery of test we examined in the paper to our case study in order to ascertain if the hospital type effect has to be considered endogenous, i.e. codetermined with the hospital choice equation. This is of paramount importance to assess whether the inference on hospital quality is of some statistical reliability.

The battery of exogeneity tests presented in the bottom part of the table provides conflicting indications at a first sight. The CM1, LM1, LR and RHO tests

[11] Notice that our final specifications do not contain any interaction term between the hospital type and the other explanatory variables, as such terms prooved not to be significant.

[12] The role played by financial incentives on clinical practices has been the subject of the long lasting literature on Physician Demand Induction (see Cromwell and Mitchell 1986 and McGuire 2000).

Table 2: Variable description.

Variable	Description
Cesarean	=1 if woman delivers with cesarean section; =0 otherwise
Private	=1 if woman delivers in a private hospital; =0 otherwise
	Risk factors
No problem	=1 if the woman self-reports not having suffered any health problem during her pregnancy; =0 otherwise
Diabetes	=1 if the woman self-reports having suffered from diabetes during her pregnancy; =0 otherwise
Gestosis	=1 if the woman self-reports having suffered from "gestosi" during her pregnancy; =0 otherwise
Hyperten	=1 if the woman self-reports having suffered from blood hypertension during her pregnancy; =0 otherwise
Twin	=1 if multiple delivery
BMI	Body Mass Index (=bodyweight/(height/100)2)
Weight gain	=1 if the woman experienced a body weight increase of more than 20 kg; =0 otherwise
Newborn weigth	weight of the newborn in kilograms
Newborn weigth sq	weight of the newborn squared
Amniocen	=1 if the woman underwent early prenatal diagnostic checks ("villi coriali" or "amniocentesi"); =0 otherwise
No. scans	number of fetal ultrasound scans done during pregnancy
Hospitalization	=1 if the woman was admitted to hospital during her pregnancy; =0 otherwise
Smoked	=1 if the woman was an abitual smoker; =0 otherwise
Age -26	=1 if woman is younger than 26; =0 otherwise
Age +36	=1 if woman is older than 36; =0 otherwise
Age	age in years
Agesq	age squared
	Socio-economic variables
Edu-high	=1 if woman holds an high education degree; =0 otherwise
Edu-low	=1 if woman holds a low education degree; =0 otherwise
Edu-medium	=1 if woman holds a medium education degree; =0 otherwise
Insured	=1 if the woman is covered by private health insurance
Self-employed	=1 if the woman is self-employed; =0 otherwise
	Other controls
NW	=1 if the woman resides in a North-West region; =0 otherwise
NE	=1 if the woman resides in a North-East region; =0 otherwise
CEN	=1 if the woman resides in a Centre region; =0 otherwise
ISL	=1 if the woman resides in Sicily or Sardinia; =0 otherwise
Area-metro	=1 if the woman resides in a metropolitan area; =0 otherwise
Area-suburb	=1 if the woman resides in a metropolitan suburb; =0 otherwise
Area-small	=1 if the woman resides in a very small commune (less than 2000 inhabitants); =0 otherwise
Area-medium	=1 if the woman resides in a medium-small commune (between 2000 and 10000 inhabitants); =0 otherwise

lead to strong rejection of the hypothesis of exogeneity, while LM2, LM3 and LM4 support the opposite evidence, i.e. in favour of exogeneity of the hospital type dummy. However, our Monte Carlo evidence helps in distinguishing and interpreting these results, as the latter set of tests has been shown to exhibit finite sample distributions remarkably far from the asymptotic ones.[13] This leads us to conclude that in our case study the bivariate endogenous dummy model is the more appropriate setting for drowing some consistent inference on hospital type differences in CS utilization rates.

Table 3: Tests for the exogeneity of the hospital type dummy.

	Estimation results		
Probit model		Private	RHO
Univariate			
	Coeff.	0.2193	-
	Std. err	0.0804	-
Bivariate SURE			
	Coeff.	-	0.0989
	Std. err	-	0.0414
Bivariate endogenous dummy			
	Coeff.	1.3253	-0.5738
	Std. err	0.3475	0.1719
	Exogeneity tests		
		statistic	p value
Bivariate endogenous dummy			
	CM1	-2.2500	0.0250
	LM1	34.0205	0.0000
	LM2	0.9054	0.3413
	LM3	0.3490	0.5547
	LM4	0.3360	0.5621
	LR	5.9248	0.0149
	RHO	-3.3377	0.0008

When we switch to the bivariate endogenous dummy model we see that the hospital type effect is codetermined with the hospital choice equation. The estimated correlation coefficient is negative and significant. This means that, conditioning on the observables, the higher (the lower) the probability of choosing a private (public) hospital the lower (the higher) the probability of delivering

[13]The nominal power of these tests -evaluated with the asymptotic critical values- was found to be very low compared to the other tests. At 5% level, the power in DGP2 was almost zero for LM2,LM3, LM4 $N = 2000$, $\rho = -0.75$, while around 80% for the other tests. In DGP3 the corresponding figures where 18% for LM2, 43% for LM3, 14% for LM4, but 99% for the alternative tests. Therefore, the results given by these tests in our application can be explained by the high, in finite sample, second-type errors.

via CS. This suggests that, among the two opposing self-selection mechanisms we figured out, the "idiosyncratic preferences" seems to be dominated by the "unobserved frailty" self-selection. In other words, despite we cannot exclude that both mechanisms are at play and partially compesating each other, we have to retain the unobserved frailty mechanism as the more relevant factor driving women self-sorting into hospital types.[14] It is worth noticing that according to the bivariate seemingly unrelated probit model the estimated correlation is positive and therefore apparently coherent with an opposite interpretation of the self-selection process at work. As the bivariate seemingly unrelated probit model is actually nested in the endogenous dummy one, we are able to conclude that the former is rejected, with the coefficient of the dummy being significantly different from zero. Moreover, the consequent structure is fundamental in order to derive quality assessment implications on the two types of provider. At the extreme case in which we interpret the selection effect as entirely attributable to woman's unobservable frailty, we have to retain that the higher utilization of CS delivery we observe in private hospitals does reflect an even lower level of healthcare quality than that we were able to identify within the univariate framework.

5 Conclusions

The bivariate probit model with endogenous dummy is the appropriate inference tool in many applications where there are good "a priori" reasons to consider a dependent binary variable as simultaneously determined with a dycotomous regressor. Modeling mortality or treatment choice as a function of healthcare provider indicators is a leading example in health economics. In this paper we adopt a bivariate probit model with endogenous hospital type dummy to analyse Cesarean Section (CS) utilization variation across public and private hospitals. Joint estimation of the two equations in this model can be avoided when the dummy variable is exogenous for the equation of interest.

Different approaches for testing the hypothesis of exogeneity in this context are available. Some tests involve only univariate probit estimation, while others are based on the bivariate model simultaneous estimation. We explore the finite sample behaviour of available alternative test statistics and evaluate their relative performance through an extensive set of simulation experiments. From the results of the Monte Carlo investigation we can draw some useful guidelines for the applied econometrics practice. First, among the testing approaches based on single equation estimation, Conditional Moment test and the simplest Lagrange Multiplier based on auxiliary regression (LM1 in our notation) perform generally well, only slightly worse than the Likelihood Ratio test. This finding is encouraging and particularly relevant in the context of multivariate probit

[14] We obtain corroborating evidence for this conclusion applying the same model on a large sample coming from an Italian region (Emilia-Romagna) discharge database that comprise a full set of controls for clinical predisposing factors. Our testing procedure supports there the exogeneity of the hospital type dummy.

models involving more than two simultaneous equations, where avoiding joint estimation is even more desirable. Second, alternative computational choices in the evaluation of the LM statistic, like that suggested in Greene's textbook (LM3 in our notation), originate test statistics that are seriously under-sized, and therefore should be applied with enormous caution. The final indication we derive is that the inference on the estimated correlation coefficient is problematic, and requires a lot of sample information. Whenever the simultaneous estimation is performed, we therefore recommend that the exogeneity status is evaluated through the Likelihood Ratio test.

In our case study on an Italian sample we apply the whole battery of tests and obtain strong evidence against the hypothesis of exogeneity of hospital type dummy in the equation determining CS probability. Our results suggest that a self-selection mechanism allocating the more risky patients to public hospital is prevailing, and that, after controlling for observable and unobservable characteristics, women admitted to a private hospital show higher probabilities of delivering via CS. This is a valuable indication for the policy maker as it testifies the existence of clear overuse of such a practice in private hospitals.

References

[1] Agency for Healthcare Research and Quality. 2002. AHRQ Quality Indicators - Guide to Inpatient Quality Indicators: Quality of Care in Hospitals - Volume, Mortality, and Utilization. Rockville, MD. Revision 2.

[2] H. Benitez-Silva, M. Buchinsky, H. Chan, S. Cheidvasser and J. Rust. How Large is the Bias in Self-Reported Disabilty? *Forthcoming* Journal of Applied Econometrics....

[3] Bryson A., Cappellari, L., Lucifora, C. 2003. Does Union Membership Really Reduce Job Satisfaction?, Quaderni dell'Istituto di Economia dell'Impresa e del Lavoro n. 34. Università Cattolica del Sacro Cuore, Milano.

[4] Buchmueller, T.C., Couffinhal, A., Grignon, M., Perronin, M. 2002. Access to Physician Services: Does Supplemental Insurance Matter? Evidence from France. W. P. 9238, NBER, Cambridge.

[5] Chesher AD, Irish M. 1987. Residual Analysis in Grouped Data and Censored Normal Linear Models. Journal of Econometrics 34: 33-62.

[6] Cromwell J,. Mitchell J. 1986. Physician-induced Demand for Surgery. Journal of Health Economics 5: 293-313.

[7] Davidson R, MacKinnon JG. 1984. Convenient Specification Tests for Logit and Probit Models. Journal of Econometrics 25: 241-262.

[8] Davidson R, MacKinnon JG. 1989. Testing for Consistency Using Artificial Regressions. Econometric Theory 5: 363-384.

[9] Deadman D, MacDonald Z. 2004. Offenders as Victims of Crime?: an Investigation into the Relationship between Criminal Behaviour and Victimization. Journal of the Royal Statistical Society A, 167, Part 1: 53-67.

[10] Dubay L, Kaestner R, Waidmann T. 1999. The Impact of Malpractice Fears on Cesarean Section Rates. Journal of Health Economics 18: 491-522.

[11] Fabbri, D. and C. Monfardini, 2001, "Demand induction with a discrete distribution of patients", Working Paper n° 414, Dipartimento di Scienze Economiche, Università di Bologna.

[12] Geweke J, Gowrisankaran G, Town RJ. 2003. Bayesian Inference for Hospital Quality in a Selection Model. Econometrica 71: 1215-1238.

[13] Greene WH. 2003. Econometric Analysis, Prentice Hall.

[14] Gruber J, Owings M. 1996. Physician Financial Incentives and Cesarean Section. Rand Journal of Economics 27: 99-123.

[15] Heckman J. 1978. Dummy Endogenous Variables in a Simultaneous Equation System. Econometrica 46: 931-959.

[16] Holly A, Gardiol L, Domenighetti G, Bisig B. 1998. An Econometric Model of Health Care Utilization and Health Insurance in Switzerland. European Economic Review 42: 513-522.

[17] Kiefer NM. 1982. Testing for Dependence in Multivariate Probit Models. Biometrika 69: 161-166.

[18] Kooreman P. 1994. Estimation of Econometric Models of Some Discrete Games. Journal of Applied Econometrics 9: 255-268.

[19] Maddala GS. 1983. Limited Dependent and Qualitative Variables in Econometrics Cambridge University Press: Cambridge.

[20] Maddala GS, 1995 Specification Tests in Limited Dependent Variable Models in Maddala GS,Phillips PCB, Srinivasa TN eds. Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor CR Rao, Blackwell, Oxford UK and Cambridge USA.

[21] McGuire TG. 2000. Physician Agency. In Handbook of Health Economics, chapter 9, Culyer AJ, Newhouse JP (eds). Elsevier: Amsterdam:

[22] Newey WK. 1985. Maximum Likelihood Specification Testing and Conditional Moment Tests. Econometrica 53: 1047-1070.

[23] Pagan A, Vella F. 1989. Diagnostic Tests for Models Based on Individual Data: a Survey. Journal of Applied Econometrics 4: S29-S59.

[24] Pudney S, Shields M. 2000. Gender, Race, Pay and Promotion in the British Nursing Profession: Estimation of a Generalized Ordered Probit Model. Journal of Applied Econometrics 15: 367-399

[25] Rivers D, Vuong Q. 1988. Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models. Journal of Econometrics 39: 347-366.

[26] Skeels CL, Vella F. 1999. A Monte Carlo Investigation of Sampling Behavior of Conditional Moment Tests in Tobit and Probit Models. Journal of Econometrics 92: 275-294.

[27] Smith RJ, Blundell RW. 1986. An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply. Econometrica 54: 679-685.

[28] Smith MD, Moffat PG. 1999. Fisher's Information on the Correlation Coefficient in Bivariate Logistic Models. Australian and New Zealand Journal of Statistics 41: 315-330.

[29] Tauchen G. 1985. Diagnostic Testing and Evaluation of Maximum Likelihood Models. Journal of Econometrics 30: 415-443.

[30] Wilde J. 2000. Identification of Multiple Equation Probit Models with Endogenous Dummy Regressors. Economics Letters 69: 309-312.

[31] World Health Organization. 1985. Recommendation on Appropriate Technology for Birth. document approved in the WHO Conference at Fortaleza, Brazil, 22-26 april 1985.

APPENDIX 1

Monte Carlo Results – 5000 replications

Table 1.1
Finite sample behaviour of exogeneity test statistics - Empirical size (ρ=0)

DGP1

		CM1	LM1	LM2	LM3	LM4	LR	RHO
N=500								
	10%	0.1154	0.1172	0.0028	0.0038	0.0028	0.1022	0.1240
	5%	0.0632	0.0644	0.0004	0.0006	0.0002	0.0536	0.0740
	1%	0.0192	0.0192	0.0000	0.0000	0.0000	0.0148	0.0274
N=1000								
	10%	0.1090	0.1100	0.0034	0.0044	0.0036	0.1054	0.1156
	5%	0.0586	0.0586	0.0004	0.0008	0.0006	0.0552	0.0654
	1%	0.0112	0.0114	0.0000	0.0000	0.0000	0.0102	0.0156

DGP2

		CM1	LM1	LM2	LM3	LM4	LR	RHO
N=1000								
	10%	0.1224	0.1228	0.0000	0.0000	0.0000	0.1164	0.1848
	5%	0.0720	0.0724	0.0000	0.0000	0.0000	0.0632	0.1356
	1%	0.0178	0.0180	0.0000	0.0000	0.0000	0.0132	0.0784
N=2000								
	10%	0.1152	0.1156	0.0000	0.0000	0.0000	0.1084	0.1444
	5%	0.0576	0.0576	0.0000	0.0000	0.0000	0.0548	0.0866
	1%	0.0132	0.0134	0.0000	0.0000	0.0000	0.0120	0.0340

DGP3

		CM1	LM1	LM2	LM3	LM4	LR	RHO
N=1000								
	10%	0.1170	0.1174	0.0000	0.0004	0.0000	0.1038	0.1464
	5%	0.0596	0.0600	0.0000	0.0000	0.0000	0.0508	0.0946
	1%	0.0148	0.0150	0.0000	0.0000	0.0000	0.0112	0.0414
N=2000								
	10%	0.1076	0.1078	0.0000	0.0000	0.0000	0.1004	0.1192
	5%	0.0548	0.0552	0.0000	0.0000	0.0000	0.0480	0.0700
	1%	0.0138	0.0138	0.0000	0.0000	0.0000	0.0104	0.0232

Table 1.2
Finite sample behaviour of exogeneity test statistics - Effect on omission of multiplicative dummy on empirical size

DGP 3

		CM1	LM1	LM2	LM3	LM4	LR	RHO
N=1000								
	10%	0.1568	0.1588	0.0002	0.0004	0.0000	0.2370	0.2542
	5%	0.0902	0.0908	0.0000	0.0000	0.0000	0.1498	0.1948
	1%	0.0264	0.0264	0.0000	0.0000	0.0000	0.0564	0.1216
N=2000								
	10%	0.1950	0.1958	0.0000	0.0006	0.0000	0.2820	0.2462
	5%	0.1184	0.1186	0.0000	0.0000	0.0000	0.1886	0.1774
	1%	0.0350	0.0350	0.0000	0.0000	0.0000	0.0770	0.0972

Table 2.1
Finite sample behaviour of exogeneity test statistics – Exact power

DGP1

		CM1	LM1	LM2	LM3	LM4	LR	RHO
ρ =0.5	**N=500**							
	10%	0.8152	0.8330	0.8290	0.8098	0.8148	0.8180	0.8104
	5%	0.7086	0.7372	0.7348	0.6898	0.7052	0.7102	0.7110
	1%	0.4426	0.4784	0.4532	0.3874	0.4106	0.4168	0.4644
	N=1000							
	10%	0.9708	0.9698	0.9716	0.9680	0.9676	0.9670	0.9722
	5%	0.9450	0.9442	0.9444	0.9342	0.9374	0.9400	0.9452
	1%	0.8590	0.8468	0.8388	0.8036	0.8148	0.8180	0.8644
ρ =-0.5	**N=500**							
	10%	0.7536	0.7758	0.7852	0.7560	0.7762	0.7840	0.7600
	5%	0.6304	0.6654	0.6808	0.6356	0.6706	0.6768	0.6480
	1%	0.3370	0.3742	0.3864	0.3410	0.3748	0.3866	0.3774
	N=1000							
	10%	0.9646	0.9630	0.9630	0.9588	0.9630	0.9636	0.9638
	5%	0.9278	0.9254	0.9254	0.9156	0.9268	0.9328	0.9282
	1%	0.8280	0.8128	0.8128	0.7764	0.8080	0.8142	0.8360
ρ =0.75	**N=500**							
	10%	0.9956	0.9962	0.9968	0.9946	0.9958	0.9958	0.9962
	5%	0.9916	0.9926	0.9916	0.9834	0.9884	0.9900	0.9918
	1%	0.9464	0.9546	0.9466	0.8954	0.9194	0.9270	0.9546
	N=1000							
	10%	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5%	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1%	1.0000	0.9996	0.9996	0.9988	0.9994	0.9996	1.0000
ρ =-0.75	**N=500**							
	10%	0.9858	0.9886	0.9910	0.9818	0.9886	0.9906	0.9890
	5%	0.9646	0.9728	0.9786	0.9582	0.9726	0.9782	0.9748
	1%	0.8634	0.8818	0.8924	0.8270	0.8726	0.8980	0.9006
	N=1000							
	10%	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
	5%	0.9998	0.9998	0.9998	0.9994	0.9998	0.9998	0.9998
	1%	0.9992	0.9990	0.9992	0.9974	0.9986	0.9990	0.9990

Table 2.2
Finite sample behaviour of exogeneity test statistics – Exact power

DGP2

		CM1	LM1	LM2	LM3	LM4	LR	RHO
ρ=0.5	**N=1000**							
	10%	0.4794	0.5098	0.5102	0.4474	0.5082	0.4972	0.5006
	5%	0.3556	0.3762	0.3768	0.2948	0.3800	0.3642	0.3772
	1%	0.1568	0.1856	0.1654	0.1060	0.1616	0.1624	0.2008
	N=2000							
	10%	0.7460	0.7534	0.7640	0.7042	0.7604	0.7644	0.7766
	5%	0.6496	0.6588	0.6552	0.5794	0.6548	0.6572	0.6812
	1%	0.3782	0.4038	0.4186	0.3168	0.4192	0.4026	0.4338
ρ =-0.5	**N=1000**							
	10%	0.3204	0.3538	0.3748	0.3734	0.3690	0.3886	0.3466
	5%	0.2088	0.2256	0.2464	0.2488	0.2488	0.2724	0.2622
	1%	0.0648	0.0818	0.0908	0.0960	0.0864	0.1106	0.1624
	N=2000							
	10%	0.5884	0.5964	0.6056	0.6010	0.5974	0.6330	0.5730
	5%	0.4682	0.4784	0.4702	0.4812	0.4676	0.5178	0.4570
	1%	0.2052	0.2318	0.2518	0.2492	0.2450	0.2792	0.2506
ρ=0.75	**N=1000**							
	10%	0.8486	0.8656	0.8802	0.8068	0.8770	0.8872	0.9014
	5%	0.7654	0.7834	0.7942	0.6686	0.8004	0.8170	0.8438
	1%	0.5444	0.5854	0.5662	0.3680	0.5642	0.6072	0.6942
	N=2000							
	10%	0.9870	0.9876	0.9890	0.9780	0.9890	0.9934	0.9954
	5%	0.9752	0.9766	0.9794	0.9508	0.9792	0.9850	0.9898
	1%	0.9010	0.9146	0.9266	0.8242	0.9240	0.9418	0.9584
ρ =-0.75	**N=1000**							
	10%	0.5470	0.5762	0.6078	0.5690	0.6018	0.6588	0.6426
	5%	0.4032	0.4274	0.4722	0.4200	0.4656	0.5342	0.5558
	1%	0.1670	0.1954	0.2372	0.2008	0.2278	0.2956	0.4248
	N=2000							
	10%	0.8452	0.8506	0.8646	0.8378	0.8584	0.8906	0.8676
	5%	0.7724	0.7792	0.7900	0.7508	0.7846	0.8362	0.8090
	1%	0.4956	0.5232	0.5740	0.5050	0.5612	0.6552	0.6608

Table 2.3
Finite sample behaviour of exogeneity test statistics – Exact power

DGP3

		CM1	LM1	LM2	LM3	LM4	LR	RHO
ρ =0.5	**N=1000**							
	10%	0.5304	0.5276	0.5842	0.5056	0.5736	0.5578	0.5480
	5%	0.4140	0.4110	0.4694	0.3730	0.4622	0.4422	0.4164
	1%	0.1980	0.1828	0.2460	0.1502	0.2248	0.2146	0.1546
	N=2000							
	10%	0.8062	0.8008	0.8345	0.7728	0.8291	0.8145	0.8205
	5%	0.7156	0.6984	0.7508	0.6703	0.7440	0.7344	0.7390
	1%	0.4688	0.4534	0.5080	0.3837	0.4910	0.4904	0.5116
ρ =-0.5	**N=1000**							
	10%	0.6642	0.6598	0.6094	0.6478	0.6054	0.6506	0.6826
	5%	0.5576	0.5532	0.4714	0.5188	0.4776	0.5364	0.5512
	1%	0.3386	0.3206	0.2236	0.2598	0.2088	0.2884	0.2156
	N=2000							
	10%	0.8980	0.8932	0.8682	0.8910	0.8678	0.8940	0.9138
	5%	0.8370	0.8266	0.7766	0.8172	0.7770	0.8284	0.8608
	1%	0.6460	0.6324	0.5044	0.5730	0.5002	0.6168	0.6890
ρ =0.75	**N=1000**							
	10%	0.8386	0.8358	0.8900	0.7914	0.8832	0.8764	0.8740
	5%	0.7540	0.7522	0.8188	0.6778	0.8088	0.8078	0.8012
	1%	0.5236	0.4994	0.6214	0.3898	0.5862	0.5928	0.5216
	N=2000							
	10%	0.9866	0.9862	0.9924	0.9788	0.9914	0.9910	0.9906
	5%	0.9746	0.9708	0.9846	0.9544	0.9824	0.9826	0.9832
	1%	0.8985	0.8937	0.9343	0.8215	0.9275	0.9341	0.9424
ρ =-0.75	**N=1000**							
	10%	0.9666	0.9664	0.9486	0.9562	0.9434	0.9610	0.9750
	5%	0.9402	0.9396	0.8926	0.9118	0.8914	0.9308	0.9468
	1%	0.8482	0.8358	0.7004	0.7420	0.6738	0.8120	0.7818
	N=2000							
	10%	0.9998	0.9998	0.9998	0.9998	0.9998	0.9994	1.0000
	5%	0.9992	0.9992	0.9978	0.9990	0.9978	0.9994	0.9998
	1%	0.9920	0.9918	0.9708	0.9818	0.9678	0.9914	0.9972

Table 3.1
Finite sample inference on correlation and dummy coefficients

DGP1

	Empirical distribution of MLE $\hat{\rho}$ (for different true values of ρ)			**Empirical distribution of MLE $\hat{\delta}_2$ (true value $\delta_2 = 1$)**		
N=500						
ρ	*mean*	*st.dev*	*m(st.dev)*	*mean*	*st.dev*	*m(st.dev)*
0	0.0067	0.2149	0.2114	0.9904	0.3226	0.3209
0.25	0.2574	0.2025	0.1971	0.9910	0.3319	0.3271
-0.25	-0.2442	0.2099	0.2048	0.9922	0.3066	0.3043
0.5	0.5059	0.1672	0.1632	0.9983	0.3274	0.3232
-0.5	-0.4947	0.1850	0.1766	0.9923	0.2858	0.2769
0.75	0.7564	0.1112	0.1069	1.0029	0.3061	0.3035
-0.75	-0.7480	0.1300	0.1235	0.9902	0.2383	0.2363
N=1000						
ρ						
0	0.0023	0.1532	0.1507	0.9989	0.2276	0.2277
0.25	0.2478	0.1421	0.1412	1.0050	0.2314	0.2326
-0.25	-0.2486	0.1462	0.1458	0.9981	0.2141	0.2155
0.5	0.4987	0.1196	0.1165	1.0038	0.2306	0.2291
-0.5	-0.4993	0.1273	0.1250	1.0032	0.1866	0.1858
0.75	0.7511	0.0783	0.0763	1.0040	0.2180	0.2142
-0.75	-0.7535	0.0890	0.0863	1.0014	0.1667	0.1650

Table 3.2
Finite sample inference on correlation and dummy coefficients

DGP2

	Empirical distribution of MLE $\hat{\rho}$ (for different true values of ρ)			**Empirical distribution of MLE $\hat{\delta}_2$ (true value $\delta_2 = 1.5$)**		
N=1000						
ρ	*mean*	*st.err.*	*mean(st.err)*	*mean*	*st.dev*	*m(st.dev)*
0	-0.0185	0.3430	0.3144	1.5401	0.6739	0.6339
0.25	0.2194	0.3126	0.2867	1.5792	0.6999	0.6737
-0.25	-0.2485	0.3400	0.3143	1.5008	0.6272	0.5878
0.5	0.4712	0.2520	0.2291	1.5945	0.7238	0.6941
-0.5	-0.4622	0.3137	0.2869	1.4452	0.5773	0.5392
0.75	0.7341	0.1556	0.1402	1.5225	0.6643	0.6641
-0.75	-0.6831	0.2690	0.2299	1.3793	0.5588	0.4923
N=2000						
ρ						
0	-0.0080	0.2416	0.2334	1.5226	0.4853	0.4685
0.25	0.2323	0.2180	0.2083	1.5464	0.5076	0.4891
-0.25	-0.2543	0.2470	0.2375	1.5091	0.4549	0.4362
0.5	0.4863	0.1668	0.1606	1.5476	0.5002	0.4924
-0.5	-0.4861	0.2289	0.2146	1.4803	0.4197	0.3943
0.75	0.7461	0.0986	0.0946	1.4816	0.4496	0.4608
-0.75	-0.7162	0.1879	0.1629	1.4403	0.3894	0.3453

Table 3.3
Finite sample inference on correlation and dummy coefficients

DGP3

	Empirical distribution of MLE $\hat{\rho}$ (for different true values of ρ)			**Empirical distribution of MLE $\hat{\delta}_2$ (true value $\delta_2 = -0.7$)**		
N=1000						
ρ	*mean*	*st.dev*	*m(st.dev)*	*mean*	*st.dev*	*m(st.dev)*
0	0.0076	0.2758	0.2658	-0.6751	0.6018	0.5744
0.25	0.2424	0.2780	0.2649	-0.6432	0.6718	0.6365
-0.25	-0.2372	0.2517	0.2420	-0.6899	0.5286	0.5114
0.5	0.4792	0.2512	0.2364	-0.6159	0.7186	0.6964
-0.5	-0.4847	0.2048	0.1946	-0.7056	0.4687	0.4465
0.75	0.7188	0.1859	0.1762	-0.6037	0.7244	0.7412
-0.75	-0.7420	0.1269	0.1208	-0.7078	0.3879	0.3680
N=2000						
ρ						
0	0.0027	0.1937	0.1918	-0.6838	0.4145	0.4056
0.25	0.2457	0.1966	0.1913	-0.6725	0.4646	0.4495
-0.25	-0.2448	0.1757	0.1722	-0.6872	0.3687	0.3588
0.5	0.4924	0.1753	0.1688	-0.6609	0.5093	0.4899
-0.5	-0.4939	0.1398	0.1358	-0.6947	0.3193	0.3100
0.75	0.7421	0.1264	0.1219	-0.6712	0.5167	0.5112
-0.75	-0.7473	0.0851	0.0834	-0.6989	0.2607	0.2543

Table 4
Empirical second-type error of the null hypothesis: $H_0 : \delta_2 = 0$

	0	**0.25**	**-0.25**	**0.5**	**-0.5**	**0.75**	**-0.75**
DGP1($\delta_2 = 1$)							
N=500							
10%	0.0836	0.0748	0.0760	0.0514	0.0574	0.0186	0.0252
5%	0.1474	0.1380	0.1284	0.1046	0.0994	0.0454	0.0454
1%	0.3258	0.3276	0.2794	0.2894	0.2078	0.1774	0.1198
N=1000							
10%	0.0028	0.0020	0.0022	0.0012	0.0018	0.0002	0.0008
5%	0.0050	0.0062	0.0048	0.0030	0.0036	0.0006	0.0012
1%	0.0362	0.0314	0.0302	0.0150	0.0168	0.0034	0.0051

DGP2 ($\delta_2 = 1.5$)							
N=1000							
10%	0.2762	0.2658	0.2700	0.2234	0.2644	0.1556	0.2480
5%	0.3912	0.3950	0.3670	0.3778	0.3452	0.3152	0.2940
1%	0.5968	0.6548	0.5356	0.7030	0.4818	0.7296	0.4002
N=2000							
10%	0.0616	0.0404	0.0670	0.0176	0.0802	0.0072	0.0956
5%	0.1164	0.0966	0.1196	0.0566	0.1208	0.0216	0.1300
1%	0.2978	0.2860	0.2768	0.2306	0.2466	0.1350	0.2140

DGP3 ($\delta_2 = -0.7$)							
N=1000							
10%	0.6922	0.7166	0.6478	0.7064	0.5448	0.7076	0.3570
5%	0.7862	0.7972	0.7622	0.7912	0.6884	0.7838	0.5216
1%	0.9072	0.8992	0.9172	0.8878	0.9034	0.8788	0.8196
N=2000							
10%	0.4852	0.5556	0.3896	0.5820	0.2494	0.5760	0.0910
5%	0.6144	0.6668	0.5400	0.6762	0.3756	0.6644	0.1632
1%	0.8164	0.8300	0.7860	0.8332	0.6738	0.8012	0.4238

Table 5
Effect of misspecification of discrete choice model on parameter estimates.

DGP1	**true value**	**univ.**	**biv. sure**	**biv. endog**
Eq: y2		coeff (se)	coeff (se)	coeff (se)
z	0.5	0.2802 (0.049)	1.2349 (0.034)	0.5550 (0.060)
y1	**1**	**1.6487 (0.058)**	**-**	**0.9461 (0.099)**
y1*z	1	0.8671 (0.074)	-	0.8485 (0.076)
cons	-0.5	-0.8955 (0.048)	0.2658 (0.021)	-0.4481 (0.071)
rho	**0.5**	**-**	**0.7810 (0.018)**	**0.5003 (0.051)**

DGP2	**true value**	**univ.**	**biv. sure**	**biv. endog**
Eq: y2		coeff (se)	coeff (se)	coeff (se)
z	-1	-1.0879 (0.031)	-0.8935 (0.024)	-0.9757 (0.044)
y1	**1.5**	**2.4365 (0.151)**	**-**	**1.2666 (0.317)**
y1*z	-1	-1.2302 (0.168)	-	-0.8884 (0.192)
cons	-0.5	-0.5944 (0.024)	-0.3754 (0.020)	-0.4680 (0.038)
rho	0.5	-	0.7945 (0.022)	0.5268 (0.103)

DGP3	**true value**	**univ.**	**biv. sure**	**biv. endog**
Eq: y2		coeff (se)	coeff (se)	coeff (se)
z	-1	-1.0384 (0.046)	-1.2238 (0.045)	-0.9018 (0.047)
y1	**-0.7**	**0.3664 (0.189)**	**-**	**-1.0420 (0.319)**
y1*z	-0.6	-1.0486 (0.176)	-	-0.6357 (0.194)
Cons	1.9	1.7822 (0.046)	1.8022 (0.048)	1.8087 (0.043)
Rho	0.5	-	-0.2686 (0.039)	0.6662 (0.096)

APPENDIX 2

Table 6
Descriptive statistics

	Full sample		Public hospital admissions (PRIV==0)		Private hospital admissions (PRIV==1)	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Priv	0.086	0.281	0.000		1.000	
No problem	0.373	0.484	0.370	0.483	0.398	0.490
Diabetes	0.019	0.138	0.019	0.137	0.023	0.149
Gestosis	0.036	0.187	0.038	0.190	0.023	0.149
Hyperten	0.044	0.206	0.046	0.208	0.031	0.173
Twin	0.017	0.129	0.016	0.127	0.021	0.142
BMI	22.73	3.471	22.73	3.487	22.68	3.301
Weight gain	0.043	0.203	0.043	0.203	0.045	0.208
Newborn weigth	3.264	0.511	3.268	0.516	3.226	0.447
No. scans	5.332	2.313	5.313	2.304	5.526	2.396
Amniocen	0.235	0.424	0.231	0.421	0.283	0.451
Hospitalization	0.549	0.498	0.547	0.498	0.571	0.495
Smoked	0.245	0.430	0.245	0.430	0.246	0.431
age	32.63	5.075	32.62	5.040	32.74	5.435
Age -26	0.111	0.314	0.110	0.312	0.125	0.331
Age +36	0.280	0.449	0.276	0.447	0.320	0.467
Edu-high	0.376	0.484	0.382	0.486	0.308	0.462
Edu-medium	0.463	0.499	0.459	0.498	0.505	0.500
Edu-low	0.102	0.302	0.097	0.296	0.152	0.359
Insured	0.156	0.363	0.155	0.362	0.164	0.371
Self-employed	0.442	0.497	0.433	0.496	0.530	0.500
NW	0.170	0.376	0.178	0.383	0.084	0.278
NE	0.207	0.405	0.215	0.411	0.119	0.324
CEN	0.161	0.368	0.165	0.371	0.125	0.331
ISL	0.130	0.336	0.127	0.333	0.156	0.363
Area-metro	0.079	0.270	0.074	0.262	0.131	0.338
Area-suburb	0.094	0.292	0.090	0.286	0.140	0.347
Area-small	0.185	0.388	0.194	0.396	0.084	0.278
Area-medium	0.284	0.451	0.285	0.452	0.273	0.446

Table 7
Full estimation results

	Univariate		Bivariate SURE		Bivariate Endogenous	
	Coeff.	*Std.err.*	Coeff.	*Std.err.*	Coeff.	*Std.err.*
Private	0.2193	*(0.0804)*			1.3253	*(0.3475)*
No problem	-0.1673	*(0.0689)*	-0.1673	*(0.0688)*	-0.1595	*(0.0679)*
Diabetes	0.1539	*(0.1729)*	0.1591	*(0.1740)*	0.1228	*(0.1642)*
Gestosis	0.2353	*(0.1254)*	0.2264	*(0.1255)*	0.2708	*(0.1223)*
Hyperten	0.1426	*(0.1107)*	0.1394	*(0.1110)*	0.1511	*(0.1065)*
Twin	0.4312	*(0.1634)*	0.4385	*(0.1643)*	0.3672	*(0.1600)*
BMI	0.0267	*(0.0069)*	0.0266	*(0.0069)*	0.0258	*(0.0067)*
Weight gain	0.1953	*(0.1135)*	0.1954	*(0.1133)*	0.1810	*(0.1089)*
Newborn weigth	-1.1295	*(0.2491)*	-1.1097	*(0.2488)*	-1.1677	*(0.2464)*
Newborn weigth sq	0.1288	*(0.0379)*	0.1253	*(0.0378)*	0.1386	*(0.0377)*
No. scans	0.0373	*(0.0099)*	0.0375	*(0.0099)*	0.0345	*(0.0098)*
Amniocen	0.0939	*(0.0569)*	0.1016	*(0.0568)*	0.0553	*(0.0573)*
Hospitalization	0.1450	*(0.0641)*	0.1452	*(0.0641)*	0.1377	*(0.0630)*
Smoked	0.0150	*(0.0541)*	0.0161	*(0.0541)*	0.0167	*(0.0521)*
Age -26	-0.0132	*(0.0683)*	-0.0079	*(0.0683)*	-0.0310	*(0.0660)*
Age +36	0.0695	*(0.0597)*	0.0702	*(0.0597)*	0.0635	*(0.0577)*
Edu-high	0.1885	*(0.1245)*	0.2143	*(0.1243)*	0.0743	*(0.1252)*
Edu-medium	0.1603	*(0.1068)*	0.1762	*(0.1067)*	0.0896	*(0.1059)*
Edu-low	0.1703	*(0.1069)*	0.1810	*(0.1069)*	0.1214	*(0.1053)*
Insured	0.1227	*(0.0654)*	0.1268	*(0.0654)*	0.1008	*(0.0643)*
NW	-0.1950	*(0.0698)*	-0.2191	*(0.0692)*	-0.0854	*(0.0749)*
NE	-0.2795	*(0.0672)*	-0.3021	*(0.0666)*	-0.1715	*(0.0735)*
CEN	-0.1829	*(0.0724)*	-0.2032	*(0.0719)*	-0.0882	*(0.0761)*
ISL	-0.0467	*(0.0744)*	-0.0562	*(0.0742)*	-0.0018	*(0.0748)*
y_2	0.0274	*(0.0573)*	0.0268	*(0.0572)*	0.0297	*(0.0551)*
y_3	0.0663	*(0.0567)*	0.0658	*(0.0566)*	0.0616	*(0.0546)*
Constant	0.7160	*(0.4589)*	0.7072	*(0.4588)*	0.7258	*(0.4504)*
No problem	-0.0333	*(0.0976)*	-0.0289	*(0.0976)*	-0.0599	*(0.0958)*
Diabetes	0.1567	*(0.2529)*	0.1602	*(0.2522)*	0.1350	*(0.2472)*
Gestosis	-0.3062	*(0.2307)*	-0.2968	*(0.2295)*	-0.3426	*(0.2308)*
Hyperten	-0.1474	*(0.1728)*	-0.1535	*(0.1722)*	-0.1135	*(0.1739)*
Twin	0.2331	*(0.2201)*	0.2380	*(0.2196)*	0.1959	*(0.2196)*
Age	-0.1073	*(0.0567)*	-0.1041	*(0.0567)*	-0.1231	*(0.0543)*
Agesq	0.0016	*(0.0009)*	0.0015	*(0.0009)*	0.0018	*(0.0009)*
Hospitalization	0.0415	*(0.0920)*	0.0405	*(0.0921)*	0.0432	*(0.0895)*
No. scans	0.0124	*(0.0136)*	0.0123	*(0.0136)*	0.0106	*(0.0132)*
Amniocen	0.1920	*(0.0781)*	0.1946	*(0.0781)*	0.1684	*(0.0792)*
Newborn weigth	0.8795	*(0.5087)*	0.8931	*(0.5098)*	0.7967	*(0.4959)*
Newborn weigth sq	-0.1573	*(0.0795)*	-0.1587	*(0.0796)*	-0.1493	*(0.0775)*
Self-employed	0.1084	*(0.0698)*	0.1044	*(0.0698)*	0.1222	*(0.0670)*
Edu-high	0.8749	*(0.1957)*	0.8680	*(0.1950)*	0.8842	*(0.1963)*
Edu-medium	0.6527	*(0.1836)*	0.6483	*(0.1829)*	0.6508	*(0.1841)*
Edu-low	0.5048	*(0.1855)*	0.5043	*(0.1849)*	0.4819	*(0.1862)*
Insured	0.1735	*(0.0906)*	0.1721	*(0.0905)*	0.1772	*(0.0886)*
NW	-0.7267	*(0.1138)*	-0.7222	*(0.1135)*	-0.7309	*(0.1132)*
NE	-0.6003	*(0.1039)*	-0.6014	*(0.1038)*	-0.5730	*(0.1049)*
CEN	-0.6177	*(0.1031)*	-0.6133	*(0.1030)*	-0.6209	*(0.1023)*
ISL	-0.2605	*(0.0968)*	-0.2602	*(0.0966)*	-0.2515	*(0.0962)*
Area-metro	0.3310	*(0.1114)*	0.3188	*(0.1112)*	0.3697	*(0.1064)*
Area-suburb	0.1465	*(0.1125)*	0.1398	*(0.1123)*	0.1734	*(0.1082)*

Area-small	-0.4163	*(0.1131)*	-0.4142	*(0.1129)*	-0.3988	*(0.1110)*
Area-medium	-0.0591	*(0.0796)*	-0.0621	*(0.0793)*	-0.0329	*(0.0786)*
Constant	-1.2256	*(1.2127)*	-1.2951	*(1.2137)*	-0.8110	*(1.1804)*
RHO			0.0989	*(0.0414)*	-0.5738	*(0.1719)*