Marginal Effects in the Bivariate Probit Model

by

William H. Greene

EC-96-11

June 1996

Marginal Effects in the Bivariate Probit Model

William H. Greene
Department of Economics
Stern School of Business
New York University
44 West Fourth St.
New York, NY, 11803
June 26, 1996

Abstract

This paper derives the marginal effects for a conditional mean function in the bivariate probit model. A general expression is given for a model which allows for sample selectivity and heteroscedasticity. The computations are illustrated using microeconomic data from a study on credit scoring.

Keywords: Marginal effects; Bivariate probit

JEL Classification: C13; C25; C35

1. Introduction

Bivariate probit models have been automated by several widely used computer programs, including TSP, Gauss, Stata, and LIMDEP, so estimation of parameters is relatively straightforward. A number of applications appear in the econometrics literature.[1] But, the computation of marginal effects, while common in other settings has been largely omitted from studies involving the bivariate probit model. Boyes, et al. (1989), for example, present only the coefficient estimates, while Van de Ven and Van Praag (1981) qualitatively examine a conditional mean (expected fraction of responses) in four different strata of their data. Analytical derivation of marginal effects, while relatively straightforward, as shown below, remains to be documented. This note will discuss direct methods for obtaining marginal effects. An example based on credit scoring data is given in the last section.

2. The Bivariate Probit Model

There are several variants of the bivariate probit model in the received literature (see, e.g., Boyes et al. (1989), Greene (1996), Maddala (1983), Meng and Schmidt (1985), Poirier (1980) and Wynand

[1]See Maddala (1983) for mention of a few.

and Van Praag (1981)). Most can be obtained as variants of the canonical model,

$$\left.\begin{aligned} y_1^* &= \beta_1'\mathbf{x}_1 + \epsilon_1,\ y_1 = \text{sgn}(y_1^*) \\ y_2^* &= \beta_2'\mathbf{x}_2 + \epsilon_2,\ y_2 = \text{sgn}(y_2^*) \end{aligned}\right\} \ [\epsilon_1,\epsilon_2] \sim \text{BVN}[0,0,1,1,\rho].$$

Note that this is a simple bivariate extension of the univariate probit model which has provided a staple of applied research for several decades. One common variant of the model results from a form of sample selection. In these settings (see, e.g., Greene (1992), Boyes, Hoffman, and Low (1989)), $[y_1,\mathbf{x}_1]$ are only observed when y_2 equals one. It has become increasingly common to accommodate the heteroscedasticity that is prevalent in microeconomic data in limited dependent variable models. For present purposes, the bivariate probit model can easily be extended to allow this by respecifying the disturbances, so that $\epsilon_j = \exp(\gamma_j'\mathbf{z}_j)u_j$, where $[u_1,u_2]$ have the previously specified bivariate standard normal joint distribution. The two variances and the covariance then become functions of the covariates, though the correlation coefficient remains ρ.

Estimation of the model by maximum likelihood is straightforward. The four cell probabilities may be written as follows, where, for convenience, we omit the observation subscripts:

$$P(y_1,y_2 \mid \mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2) = B[q_1a_1,\ q_2a_2,\ q_1q_2\rho],\ y_j = 0,\ 1 \text{ for } j = 1,2,$$

where

$$q_j = 2y_j - 1,$$

$$a_j = \beta_j'\mathbf{x}_j \,/\, \exp(\gamma_j'\mathbf{z}_j),$$

and $B(\cdot)$ denotes the bivariate normal CDF. The log-likelihood to be maximized is, then, just $\sum_i \ln P(y_{1i},y_{2i})$. In the sample selection model, the "0,0" and "1,0" cells are combined in the simple univariate probability, $\text{Prob}[y_2 = 0] = \Phi(-a_2)$, where $\Phi(\cdot)$ denotes the univariate normal CDF, and the other two cells in the joint distribution are the same.[2]

3. Marginal Effects

Once parameter estimates are obtained, a natural next step is to consider the marginal effects of the covariates in the conditional distributions. However, there is an ambiguity in this setting in that there are several conditional quantities that one might examine:

The unconditional distributions are given by the univariate normal distributions: $P(y_j) = \Phi(q_ja_j)$. The regression functions are obtained by setting q_j to one;

$$\begin{aligned} E[y_j \mid \mathbf{x}_j,\mathbf{z}_j] &= \text{Prob}[y_j = 1 \mid \mathbf{x}_j,\mathbf{z}_j] \\ &= \Phi(a_j). \end{aligned}$$

Tools for analyzing these distributions and the marginal effects in the regressions can be found in most modern econometrics textbooks, e.g., Greene (1997) or Davidson and MacKinnon (1993).

The log-likelihood for the bivariate probit model specifies four cell probabilities from the *joint*

---

[2]The log-likelihood for the selection model is derived in Wynand and Van Praag (1981) and Boyes et al. (1989).

distribution of y_1 and y_2, $P(y_1,y_2 \mid \mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2)$, from which one might derive elasticities or derivatives,

$$\mathbf{d}(y_1,y_2,\mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2) = \partial P(y_1,y_2 \mid \mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2)/\partial \mathbf{w},$$

where $\mathbf{w}$ denotes any (or all) of the covariate vectors. Other useful quantities would be the the gradients of the *conditional probabilities*,

$$\delta_1(y_1 \mid y_2,\mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2) = \partial P[y_1 \mid y_2,\mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2)/\partial \mathbf{w}$$

and, likewise for $\boldsymbol{\delta}_2(y_2 \mid y_1,\mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2)$. These vectors can also be specified for any of the four cells in the joint distribution, e.g., $\delta_i(1 \mid 1,\mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2) = \partial \text{Prob}[y_1=1 \mid y_2=1,\mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2)/\partial \mathbf{w}$. This vector has the virtue that these particular derivatives are the slopes of the conditional mean function, since

$$\text{Prob}[y_1=1 \mid y_2,\mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2] = E[y_1 \mid y_2,\mathbf{x}_1,\mathbf{x}_2,\mathbf{z}_1,\mathbf{z}_2],$$

for $y_2 = 0$ or 1. Note, however, that the log-likelihood specifies the unconditional probabilities. The conditional mean function for the bivariate probit model is

$$\begin{aligned} E[y_1 \mid y_2,x_1,x_2,z_1,z_2] &= Prob[y_1=1 \mid y_2,x_1,x_2,z_1,z_2] \\ &= \frac{Prob[y_1=1,y_2,x_1,x_2,z_1,z_2]}{Prob[y_2 \mid x_2,z_2]} \\ &= \frac{BVN[a_1,\ (2y_2-1)a_2),\ (2y_2-1)\rho]}{\Phi((2y_2-1)a_2)} \\ &= \frac{BVN[a_1,\ q_2a_2,\ q_2\rho]}{\Phi(q_2a_2)} \\ &= \frac{BVN[c_1,\ c_2,\ \rho^*]}{\Phi(c_2)}. \end{aligned}$$

We have written the functional form directly in terms of y_2 to highlight the dependence of the conditional mean function on the conditioning variable.

Derivatives of the various functions shown above give the desired marginal effects. For the bivariate distribution, these will involve the quantity

$$\begin{aligned} g_1(c_1,c_2,\rho^*) &= \frac{\partial B(c_1,c_2,\rho^*)}{\partial c_1} \\ &= \phi(c_1)\Phi\left(\frac{c_2-\rho^* c_1}{\sqrt{1-\rho^{*2}}}\right) \end{aligned}$$

(and likewise for $g_2(c_1,c_2,\rho)$). Let $\mathbf{w} = [\mathbf{x}_1 \cup \mathbf{x}_2 \cup \mathbf{z}_1 \cup \mathbf{z}_2]$ and define β_1^* with nonzero values and zeros placed appropriately so that $\beta_1^{*\prime}\mathbf{w} = \beta_1'\mathbf{x}_1$, and β_2^*, γ_1^* and γ_2^* likewise. In order to simplify the differentiation, let $c_j = q_j a_j$. Then, using standard results, we obtain the vector of marginal effects:

$$\frac{\partial E[y_1 \mid y_2, w]}{\partial w} = \left[\frac{g_1(c_1,c_2,\rho^*)}{\Phi(c_2)}\right]\left\{\left[\frac{1}{e^{\gamma_1^{*\prime} w}}\right]\beta_1^* - c_1\gamma_1^*\right\}$$

$$+ \left[\frac{\Phi(c_2)\, g_2(c_1,c_2,\rho^*) - B(c_1,c_2,\rho^*)\,\phi(c_2)}{[\Phi(c_2)]^2}\right] q_2 \left\{\left[\frac{1}{e^{\gamma_2^{*\prime} w}}\right]\beta_2^* - c_2\gamma_2^*\right\}.$$

The four parts contained in curled brackets show the effects due to changes in the means of the latent variables (y_j^*) and changes in the disturbance variances. The total effect is the sum of the four parts. If either of the disturbances are homoscedastic, the corresponding terms involving γ_j are zero.

4. Application

Greene (1994) analyzes individual data on credit scoring and loan default behavior. The following is a modification of the model presented there. Data consist of 1319 observations on applicants for a major credit card. The response variables are:

Z_1 = the number of major derogatory reports in the applicant's credit history,

y_2 = 1 if the application was accepted by the credit card vendor, 0 if denied.

The 1994 study analyzes Z_1 in the context of a model model for count data which accommodates the preponderance of zeros in the observed sample. For this application, the constructed variable used is

y_1 = 1 if $Z_1 > 0$ (i.e., 1 if any major derogatory reports) and 0 else.

The sample proportion of zero values are 0.8036 for y_1, 0.2244 for y_2, and 0.1099 for (y_1,y_2). The following bivariate probit model is specifed:

$\mathbf{x}_1$ = constant, Age, Income, Average_Monthly_Credit_Card_Expense

$\mathbf{x}_2$ = constant, Age, Income, Own/Rent_Home, Self_Employed.

The last two variables in $\mathbf{x}_2$ are binary variables. Income is scaled by 10,000. Finally, in view of the evidence in Greene (1994), we also specify

$$\text{Var}[\epsilon_1] = e^{\gamma_1 \text{Average\_Expense}}$$

In fact, the number of recent major derogatory reports is overwhelmingly the predominant factor in whether a credit card application is approved or denied. The large negative estimate of ρ shown below (-.735) is thus to be expected. Table 1 below shows the parameter estimates for the bivariate probit model.

TABLE 1. Estimated Bivariate Probit Model (Maximum Likelihood).

Variable	Coefficient	Standard Error	t-ratio.	Mean of X
Log likelihood function		-1219.890		
Equation for Derogatory Reports: Mean				
Constant	-1.3289	0.17602	-7.550	
AGE	0.11440E-01	0.47379E-02	2.415	33.21
INCOME	0.21199E-02	0.28131E-01	0.075	3.365
AVGEXP	0.11647E-03	0.13159E-02	0.089	185.1
Equation for Derogatory Reports: Variance				
AVGEXP	0.55150E-03	0.15951E-02	0.346	185.1
Equation for Card Holder Status				
Constant	0.70168	0.14567	4.817	
AGE	-0.90534E-02	0.43585E-02	-2.077	33.21
INCOME	0.73701E-01	0.23548E-01	3.130	3.365
OWNRENT	0.34091	0.80421E-01	4.239	0.4405
SELFEMPL	-0.25447	0.12216	-2.083	0.6899E-01
Disturbance Correlation				
ρ	-0.73482	0.38617E-01	-19.028	

The various parts of the marginal effects are shown in Table 2. The values reported are percentage changes. That is, the table displays $100/E[\ldots] \times \partial E[y_1 \mid y_2,\mathbf{w}]/\partial\mathbf{w}$. Estimates are computed at the means of the variables. Some authors, e.g., Hensher and Johnson (1981), advocate averaging the individual sample observations on the marginal effects instead. Our sample is relatively large, so the difference from this alternative approach would be small. Two of the effects are for binary variables. A more accurate approximation might be obtained by taking the discrete difference of the probabilities computed with these variables set to the values 1 and 0, respectively. We have found generally that this approach produces differences of no more than second order, especially in large samples, though in some circumstances, the effects could be noticeable. Since the second equation specifies a homoscedastic disturbance, the column of the table for $\mathbf{z}_2$ is omitted.

The estimated conditional mean is 0.098934, while the sample proportion $(y_1=1,y_2=1)$ is 0.08188. Table 2 lists the decomposition of the changes in this estimated proportion given changes in the model covariates.

TABLE 2. Components of Marginal Effects for E[y1 | y2 = 1,w] (% Change)

Variable	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{z}_1$	Total	Std.Error
AGE	2.8858	-1.1135	0.0000	1.7723	1.0561
INCOME	0.5348	9.0644	0.0000	9.5992	5.9814
AVGEXP	0.0294	0.0000	0.1280	0.1574	0.07084
OWNRENT	0.0000	41.9277	0.0000	41.9277	10.7892
SELFEMPL	0.0000	-31.2973	0.0000	-31.2973	15.2063

Finally, the last two columns of Table 2 present the total marginal effects, with estimated standard errors. The standard errors are computed using the delta method (Greene (1997)),

$$\text{Est.Asy.Var}[\hat{\delta}_1(y_1 \mid y_2=1,\mathbf{w})] \;=\; \mathbf{G} \times \text{Est.Asy.Var}[\hat{\boldsymbol{\theta}}] \times \mathbf{G}'$$

where $\hat{\boldsymbol{\theta}}$ is the set of maximum likelihood estimates of the bivariate probit model, including ρ, as shown in Table 1 and $\mathbf{G}$ is the estimate of the matrix of partial derivatives, $\partial\delta_1/\partial\boldsymbol{\theta}'$. The derivatives of the marginal effects with respect to the parameters are exceeedingly complicated. We used numerical approximations instead of analytic expressions.

5. Conclusions

This note has suggested an expression for the rates of change of the conditional mean function in several variants of the bivariate probit model which, apparently, has not appeared heretofore. As in many such models which involve multiple equations, the both the magnitudes and the signs of the simple coefficients in the model can be misleading.

REFERENCES

Boyes, W., D. Hoffman, and S. Low, "An Econometric Analysis of the Bank Credit Scoring Problem," *Journal of Econometrics*, 40, 1989, pp. 3-14.

Davidson, J., and J. MacKinnon, *Estimation and Inference in Econometrics*, New York: Oxford University Press, 1993.

Greene, W., "A Statistical Model for Credit Scoring," Department of Economics, Stern School of Business, New York University, Working Paper #EC-92-29, 1993.

Greene, W., "Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models," Department of Economics, Stern School of Business, New York University, Working Paper #EC-94-10, 1994.

Greene, W., *LIMDEP, Version 7.0, User's Manual*, Bellport, New York: Econometric Software, Inc., 1996.

Greene, W., *Econometric Analysis, 3rd Edition*, Englewood Cliffs, New Jersey: Prentice Hall, 1997 (forthcoming).

Hensher, D., and L. Johnson, *Applied Discrete Choice Modelling*, New York: John Wiley and Sons, 1981.

Maddala, G., *Limited Dependent and Qualitative Variables in Econometrics*, New York: Cambridge University Press, 1983.

McFadden, D. and P. Ruud, "Estimation by Simulation," *The Review of Economics and Statistics*, 76, 1994, pp. 591-608.

Meng, C., and P. Schmidt, "On the Cost of Partial Observability in the Bivariate Probit Model," *International Economic Review*, 26, 1, 1985, pp. 71-86.

Poirier, D., "Partial Observability in Bivariate Probit Models," *Journal of Econometrics*, 12, 1980, pp. 209-217.

Van de Ven, W, and B. Van Praag, "The Demand For Deductibles in Private Health Insurance: A Probit Model with Sample Selection," *Journal of Econometrics*, 17, 1981, pp. 229-252.

DEPARTMENT OF ECONOMICS WORKING PAPER SERIES 1996

Number	Author(s), Title
EC-96-01	William Easterly and Holger C. Wolf, "The Wild Ride of the Ruble," February 1996.
EC-96-02	Atish R. Ghosh and Holger C. Wolf, "On the Mark(s): Optimum Currency Areas in Germany," February 1996.
EC-96-03	Holger C. Wolf, "Savings in Eastern Europe," March 1996.
EC-96-04	Holger C. Wolf, "Currency Baskets as International Units of Account," March 1996.
EC-96-05	Gordon R. Green, "Market Institutions and Allocations," March 1996.
EC-96-06	Lawrence J. White, "U.S. Public Policy Toward Network Industries," March 1996.
EC-96-07	Nicholas Economides and Lawrence J. White, "The Inefficiency of the ECPR Yet Again: A Reply to Larson," March 1996.
EC-96-08	Barbara G. Katz and Joel Owen, "The Investment Choices of Voucher Holders and Their Impact on Privatized Firm Performance," April 1996.
EC-96-09	Martin D. D. Evans, "Index-Linked Debt and the Real Term Structure: New Estimates and Implications from the U.K. Bond Market," March 1996.
EC-96-10	Asim Ansari, Nicholas Economides, and Joel Steckel, "The Max-Min-Min Principle of Product Differentiation," May 1996.
EC-96-11	Greene, William H., "Marginal Effects in the Bivariate Probit Model," June 1996.

DEPARTMENT OF ECONOMICS WORKING PAPER SERIES 1995

Number	Author(s), Title
EC-95-01	Simon Gilchrist and Charles P. Himmelberg, "Evidence on the Role of Cash Flow for Investment," January 1995.
EC-95-02	Simon Gilchrist and Charles P. Himmelberg, "Hypothesis Testing with Two-Step GMM Estimators," January 1995.
EC-95-03	Giuseppe Lopomo, "Optimality and Robustness of the English Auction," October 1995.
EC-95-04	Nicholas Economides and Lawrence J. White, "Access and Interconnection Pricing: How Efficient is the "Efficient Component Pricing Rule"?" March 1995.
EC-95-05	Martin D. D. Evans, "Peso Problems: Their Theoretical and Empirical Implications," May 1995.
EC-95-06	William H. Greene, "Sample Selection in the Poisson Regression Model," May 1995.
EC-95-07	Chris Stefanadis, "Can Import Protection Become Export Promotion?" July 1995.
EC-95-08	Nicholas Economides and Robert A. Schwartz, "Equity Trading Practices and Market Structure: Assessing Asset Managers' Demand for Immediacy," August 1995.
EC-95-09	Nicholas Economides, "The Incentive of a Multiproduct Monopolist to Provide All Goods," August 1995.
EC-95-10	Nicholas Economides and Steven S. Wildman, "Monopolistic Competition with Two-Part Tariffs," August 1995.
EC-95-11	Nicholas Economides and Charles Himmelberg, "Critical Mass and Network Size with Application to the US FAX Market," August 1995.
EC-95-12	Nicholas Economides and Fredrick Flyer, "Technical Standards Coalitions for Network Goods," October 1995.

EC-95-13 Martin D.D. Evans, "Dividend Variability and Stock Market Swings," September 1995.

EC-95-14 Robert E. Cumby and Martin D.D. Evans, "The Term Structure of Credit Risk: Estimates and Specification Tests," September 1995.

EC-95-15 Andrew Schein, "The Variances of Wholesale and Retail Prices: Tests of a Marshallian Hypothesis," October 1995.

EC-95-16 Vittorio Grilli and Nouriel Roubini, "Liquidity Models in Open Economies: Theory and Empirical Evidence," October 1995.

EC-95-17 Vittorio Grilli and Nouriel Roubini, "Liuidity and Exchange Rates: Puzzling Evidence from the G-7 Countries, " Octover 1995.

EC-95-18 Gian Maria Milesi-Ferretti and Nouriel Roubini, "Growth Effects of Income and Consumption Taxes: Positive and Normative Analysis," October 1995.□