CER-ETH – Center of Economic Research at ETH Zurich

A note on the different interpretation of the correlation parameters in the Bivariate Probit and the Recursive Bivariate Probit

M. Filippini, W. Greene, N. Kumar, A. Martinez.Cruz

Working Paper 17/275
September 2017

Economics Working Paper Series

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

A note on the different interpretation of the correlation parameters in the Bivariate Probit and the Recursive Bivariate Probit*

Massimo Filippini[1,2],
Willian H. Greene[3],
Nilkanth Kumar[1], and
Adan L. Martinez-Cruz† [1,4]

[1] *Center of Economic Research (CER-ETH), ETH Zürich*
[2] *Università della Svizzera Italiana*
[3] *Department of Economics, NYU Stern School of Business*
[4] *Department of Economics, Centro de Investigación y Docencia Económicas (CIDE)*

Last revision: September 15, 2017

Abstract

This note makes the point that, if a Bivariate Probit (BP) model is estimated on data arising from a Recursive Bivariate Probit (RBP) process, the resulting BP correlation parameter is a weighted average of the RBP correlation parameter and the parameter associated to the endogenous binary variable. Two corollaries follow this proposition: i) a zero correlation parameter in a BP model, usually interpreted as evidence of independence between the binary variables under study, may actually mask the presence of a RBP process; and ii) the interpretation of the correlation parameter in the RBP is not the same as in the BP —i.e. the RBP correlation parameter does not necessarily reflect the correlation between the binary variables under study.

Keywords: Bivariate probit; Recursive Bivariate probit; Tetrachoric correlation; Monte Carlo simulation

JEL Classification Codes: C25

*We are grateful to CEPE for sharing the data for the purpose of this paper. All omissions and remaining errors are our responsibility.

†Department of Economics, Centro de Investigación y Docencia Económicas (CIDE), Circuito Tecnopolo Norte 117, 20313, Aguascalientes, AGS, Mexico <adan.martinez@cide.edu>

1 Introduction

In modelling two jointly determined binary choices, empirical researchers may adopt a Bivariate Probit (BP) approach. The BP is a system of two seemingly unrelated probit equations in which the correlation between the binary variables under analysis is captured by the conditional tetrachoric correlation of the error terms (Greene, 2018). A zero BP correlation parameter is usually interpreted as evidence that the binary variables can be modelled as independent of each other (e.g., Humphreys et al., 2014).

This note makes the point that a zero BP correlation parameter does not always imply independence of the binary variables under analysis. In particular, a zero correlation parameter may result from erroneously estimating a BP model on data *truly* arising from a Recursive Bivariate Probit (RBP) process. The RBP is a system of two probit equations that allows the errors terms to be correlated, and the binary dependent choice in one equation to be an endogenous regressor in the other equation.

We show that the correlation parameter in a BP model that is estimated on RBP data is a weighted average of the RBP correlation parameter and the parameter associated to the endogenous binary variable. Plausibly, this weighted average can take value zero —which current empirical practice would erroneously interpret as evidence of independence.

As a corollary, it becomes clear that the interpretation of the RBP correlation parameter does not follow the interpretation of the BP correlation parameter. That is, the RBP correlation parameter does not necessarily capture the correlation between the binary variables under analysis —i.e. once the effect of the endogenous variable is taken into account, the correlation between the errors terms is not necessarily of the same sign as the endogenous relationship.

We illustrate these implications through a series of Monte Carlo simulations and an empirical application.

2 Correlation parameter in a BP when the data follow a RBP process

Consider a *true* data-generating process that follows the Bivariate Probit with the recursive structure proposed by Maddala (1986):

$$y_{1i}^* = \boldsymbol{\beta}_1' x_{1i} + \upsilon_{1i}, \quad y_{1i} = 1 \text{ if } y_{1i}^* > 0,\ y_{1i} = 0 \text{ otherwise,} \tag{1}$$
$$y_{2i}^* = \delta y_{1i} + \boldsymbol{\beta}_2' x_{2i} + \upsilon_{2i}, \quad y_{2i} = 1 \text{ if } y_{2i}^* > 0,\ y_{2i} = 0 \text{ otherwise,} \tag{2}$$

$$[\upsilon_{1i}, \upsilon_{2i}] \sim \Phi_2[(0,0),(1,1),\zeta], \quad \zeta \in [-1, 1]$$

where i is the individual index; y_{1i}^* and y_{2i}^* are latent continuous variables for which only the binary variables y_{1i} and y_{2i} are observable; x_{1i} and x_{2i} are vectors of exogenous variables —which can possibly be identical (Wilde, 2000); and $(\upsilon_{1i}, \upsilon_{2i})'$ is a vector of error terms described by Φ_2 —a bivariate standard normal distribution with correlation ζ.[1]

Assume now that an empirical researcher estimates a Bivariate Probit that misses the recursive structure of equations (1) and (2); i.e.

$$y_{1i}^* = \boldsymbol{\beta}_1' x_{1i} + \upsilon_{1i}, \quad y_{1i} = 1 \text{ if } y_{1i}^* > 0,\ y_{1i} = 0 \text{ otherwise,} \tag{3}$$
$$y_{2i}^* = \boldsymbol{\beta}_2' x_{2i} + \varepsilon_{2i}, \quad y_{2i} = 1 \text{ if } y_{2i}^* > 0,\ y_{2i} = 0 \text{ otherwise,} \tag{4}$$

[1] A recursive structure is logically consistent with, for instance, the health production model in which individuals first engage in a healthy behaviour (y_1^*) in order to produce health (y_2^*) (see Humphreys et al., 2014).

$$[\upsilon_{1i}, \varepsilon_{2i}] \sim \Phi_2[(0,0),(1,1),\rho], \quad \rho \in [-1,1]$$

where ρ is the correlation between υ_{1i} and ε_{2i}.

If the BP defined by equations (3) and (4) is estimated on the data generated by equations (1) and (2), then the true recursive component is absorbed by the error term of equation (4) which implies that ρ is mechanically determined by ζ and δ; i.e.,

$$\begin{aligned}
\rho \equiv corr(\upsilon_{1i}, \varepsilon_{2i}) &= corr(\upsilon_{1i}, \delta y_{1i} + \upsilon_{2i}) \\
&= \frac{cov(\upsilon_{1i}, \delta y_{1i} + \upsilon_{2i})}{\sqrt{var(\upsilon_{1i}) var(\delta y_{1i} + \upsilon_{2i})}} \\
&= \frac{cov(\upsilon_{1i}, \delta y_{1i}) + cov(\upsilon_{1i}, \upsilon_{2i})}{\sqrt{var(\delta y_{1i}) + var(\upsilon_{2i}) + 2cov(\upsilon_{2i}, \delta y_{1i})}} \\
&= \frac{cov(\upsilon_{1i}, y_{1i})\delta + \zeta}{\sqrt{\delta^2 var(y_{1i}) + 1}}, \qquad (5)
\end{aligned}$$

Not surprisingly, according to equation (5) if $\delta = 0$ then $\rho = \zeta$ —i.e. in the absence of a recursive structure, the RBP collapses to the BP. Also, equation (5) shows that ρ can plausibly take value zero, depending on the signs and relative magnitude of ζ and δ —i.e. a BP model estimated on RBP data can potentially deliver a zero correlation parameter which would erroneously be interpreted as evidence of independence between y_1 and y_2. Implicit in the previous statement (and in the setting of the RBP process), no a priori restrictions are imposed on the signs of ζ and δ. For instance, ζ and δ may have opposite signs. while ζ captures the correlation between the error terms υ_1 and υ_2, it does not reflect the correlation between y_1 and y_2. Such correlation is subsumed into δ —which implies that the interpretation of the RBP correlation does not resemble the interpretation of the BP correlation parameter. Section 3 illustrates the implications from equation (5).

3 Illustration

The Monte Carlo simulations in this section are designed to illustrate how the sign of $\hat{\rho}$ depends on the signs and values of both ζ and δ —i.e. we illustrate that ρ may be estimated at zero, hiding a true recursive structure. Also, we borrow data from Blasch et al. (2017) to illustrate that, in empirical applications, ζ does not necessarily reflect the direction of the correlation between y_1 and y_2.

3.1 Monte Carlo simulations

A pseudo-population of $100,000$ individuals has been simulated according to the following recursive data-generating process:

$$y_{1i}^* = \quad -2.00 + 0.10z_{1i} + 0.90z_{2i} + \upsilon_{1i}, \quad y_{1i} = 1 \text{ if } y_{1i}^* > 0, \ y_{1i} = 0 \text{ otherwise}, \qquad (6)$$

$$y_{2i}^* = \delta y_{1i} - 1.00 + 1.20z_{1i} - 0.20z_{2i} + \upsilon_{2i}, \quad y_{2i} = 1 \text{ if } y_{2i}^* > 0, \ y_{2i} = 0 \text{ otherwise}, \qquad (7)$$

where z_1 and z_2 are two exogenous variables. Realizations of z_1 are drawn from a binomial distribution with probability of success of 0.5; and realizations of z_2 are drawn from a normal distribution with mean 2 and unitary standard deviation.

Results reported in table 1 illustrate the values of $\hat{\rho}$ that arise from erroneously estimating a BP on RBP data generated according to equations (6) and (7). Each set of results in table 1 arise from

1,000 Monte Carlo replications.[2]

Table 1: Monte Carlo simulations (1,000 replications)

	$\zeta = 0.00$, then $\rho \hat{=} 0.00$ when $\delta = 0.00$ ($\hat{\rho}$'s sign and magnitude is determined by δ)					
True parameters						
ζ	0.00	0.00	0.00	0.00	0.00	0.00
δ	2.00	1.50	0.40	0.00	−1.50	−2.00
Estimated parameters						
$\hat{\rho}$ —via a BP	0.85	0.72	0.23	0.00	−0.71	−0.82
Std. Dev.	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
$\hat{\zeta}$ —via a RBP	0.00	0.00	0.00	0.04	−0.00	0.00
Std. Dev.	(0.07)	(0.02)	(0.02)	(0.05)	(0.03)	(0.02)
$\hat{\delta}$ —via a RBP	2.00	1.48	0.39	−0.08	−1.49	−2.01
Std. Dev.	(0.11)	(0.03)	(0.05)	(0.08)	(0.04)	(0.05)
	$\zeta = 0.80$; then $\rho \hat{=} 0.00$ when δ is negative and large (≈ -1.50) ($\hat{\rho}$'s sign and magnitude is determined by both δ and ζ)					
True parameters						
ζ	0.80	0.80	0.80	0.80	0.80	0.80
δ	2.00	1.50	0.40	0.00	−1.50	−2.00
Estimated parameters						
$\hat{\rho}$ —via a BP	0.99	0.98	0.90	0.79	−0.09	−0.40
Std. Dev.	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
$\hat{\zeta}$ —via a RBP	0.80	0.79	0.81	0.79	0.81	0.80
Std. Dev.	(0.01)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)
$\hat{\delta}$ —via a RBP	1.97	1.53	0.37	0.02	−1.51	−1.98
Std. Dev.	(0.03)	(0.12)	(0.01)	(0.01)	(0.01)	(0.04)
	$\zeta = -0.80$; then $\rho \hat{=} 0.00$ when δ is positive and large ($\hat{\rho}$'s sign and magnitude is determined by both δ and ζ)					
True parameters						
ζ	−0.80	−0.80	−0.80	−0.80	−0.80	−0.80
δ	2.00	1.50	0.40	0.00	−1.50	−2.00
Estimated parameters						
$\hat{\rho}$ —via a BP	0.45	0.11	−0.63	−0.79	−0.98	−0.99
Std. Dev.	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
$\hat{\zeta}$ —via a RBP	−0.78	−0.79	−0.80	−0.79	−0.79	−0.82
Std. Dev.	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.02)
$\hat{\delta}$ —via a RBP	1.97	1.49	0.40	0.00	−1.54	−1.92
Std. Dev.	(0.02)	(0.01)	(0.04)	(0.01)	(0.09)	(0.12)

The first panel of table 1 illustrates that when no correlation between unobservables is present in the true RBP process ($\zeta = 0$),[3] then the correlation parameter in a BP model takes value zero when the parameter associated to the endogenous variable takes value zero —i.e. the sign and magnitude of $\hat{\rho}$ is determined by the sign and magnitude of δ. In the six scenarios of the first panel, ζ is assumed to be zero; from left to right δ takes values 2.00, 1.50, 0.40, 0.00, -1.50, and -2.00, respectively. Consistently with equation (5), $\hat{\rho}$ is positive when δ is positive; zero when $\delta = 0$; and negative when δ is negative. We also report $\hat{\zeta}$ and $\hat{\delta}$ to document that the correct RBP model yields estimates that reflect the true parameters.

The second panel of table 1 illustrates that when ζ is large and positive (e.g., $\zeta = 0.80$), then $\hat{\rho}$ is large and positive when δ is also positive. However, $\hat{\rho}$ can take value zero and even shift its sign if

[2]Tangentially, notice that our estimations do not require exclusion restrictions for identification purposes. See Humphreys et al. (2014) for further details on the trend of the literature that considers that exclusion restrictions are essential in estimating RBP models.

[3]This scenario is not unseen in empirical applications. For instance, Greene (1998)'s RBP yields a zero correlation parameter.

δ is negative and relatively large —in this scenario, $\delta \approx -1.50$ provokes the shift in sign. In the six scenarios of the second panel, $\zeta = 0.80$, and δ goes from 2.00 to -2.00 in a similar fashion as in the previous panel of results. Notice that $\hat{\rho} = 0.99$ when $\delta = 2.00$, and $\hat{\rho} = 0.90$ even when $\delta = 0.40$. A δ near -1.50 is needed to shift the sign of $\hat{\rho}$ —i.e. under this scenario, δ is required to be negative and large for a shift in sign to occur.

Following a similar reasoning as in the second panel, the third panel of table 1 illustrates that a large and negative correlation between unobservables in the RBP process (e.g., $\zeta = -0.80$) translates into a large and negative $\hat{\rho}$ unless δ is large and positive.

3.2 Empirical example

We borrow our empirical illustration from Blasch et al. (2017) who implement an online randomized controlled experiment in which respondents take two binary decisions. One decision is the identification, from among two alternatives, of the refrigerator with the lowest lifetime cost. Another decision consists on carrying out (or not) an investment calculation to determine the lifetime cost of the two refrigerators.

Blasch et al. (2017) hypothesize that performing an investment calculation has a positive causal effect on the probability of identifying the refrigerator with the lowest lifetime cost. Thus a RBP model is expected to yield a statistically significant positive coefficient associated to the investment calculation variable —i.e. a $\hat{\delta} > 0$.

Table 2: Selected parameters from Bivariate probit (BP) and Recursive bivariate probit (RBP) specifications estimated on data borrowed from Blasch et al. (2017).

	Number of observations: 877	
	BP	**RBP**
Investment Calculation Equation . . .		
Energy literacy index	0.0444**	0.0437**
	(0.0205)	*(0.0203)*
Investment literacy index	0.4632***	0.6643***
	(0.1317)	*(0.1361)*
Monetary treatment	0.5661***	0.5467***
	(0.1144)	*(0.1138)*
Refrigerator Choice Equation . . .		
Energy literacy index	0.0808***	0.0457**
	(0.0189)	*(0.0200)*
Monetary treatment	0.6885***	0.2971***
	(0.1074)	*(0.1151)*
Investment calculation	—	2.5436***
	—	*(0.1657)*
Correlation parameter	0.6633***	-0.8162***
	(0.0524)	*(0.1184)*

***, **, * $\Rightarrow$ Significance at 1%, 5%, 10% level.
Robust standard error in parenthesis.

Table 2 reports selected parameter estimates from a BP model and a RBP model. These models are estimated on a sample of 877 Swiss respondents analyzed by Blasch et al. (2017). In addition to the correlation estimates and the parameter associated with the investment calculation decision, we report the parameters associated to three variables of interest in Blasch et al. (2017): i) an energy literacy index; ii) an investment literacy index; and iii) a randomly allocated monetary treatment under which respondents see the yearly electricity consumption of the refrigerators in monetary units (CHF) instead of physical units (kWh).

Focusing our attention on the results from the RBP, we illustrate that ζ does not capture the correlation between y_1 (carrying out an investment calculation) and y_2 (identifying the refrigerator with the lowest lifetime cost). The positive correlation, subsumed by the positive causal relationship, is captured by δ ($\hat{\delta} = 2.54$). In this application, $\hat{\zeta} = -0.82$ —i.e. the correlation between unobserved factors is negative. These empirical results resemble the first scenario reported in the third panel of table 1 ($\zeta = -0.80$, $\delta = 2.00$). Consequently, we could expect that a BP model would deliver a ρ with the opposite sign of ζ —expectation that is confirmed ($\hat{\rho} = 0.66$).

4 Conclusions

This note makes the point that a BP delivering a zero correlation parameter may not always be interpreted as evidence of independence between the binary variables under study. It is possible that a zero correlation parameter is hiding the presence of a RBP process. Thus empirical researchers would benefit from estimating a RBP specification when a BP delivers a zero correlation parameter to check for consistency in results —a RBP specification should yield an insignificant parameter associated with the endogenous binary variable and, consequently, an identical correlation parameter than the BP. When estimating a RBP model, however, keep in mind that deciding which variable is y_1 is neither straightforward nor trivial. We should not arbitrarily choose y_1 but instead follow a careful selection process informed by a theoretical framework.

Empirical researchers should also keep in mind that the RBP correlation parameter is not interpreted in a similar fashion as the BP correlation parameter. Thus there is no need of interpreting a RBP correlation parameter as reflecting the correlation of the binary variables under analysis. Consequently, it is plausible to observe a scenario in which a BP model delivers a correlation parameter with the opposite sign yielded by a RBP model. Such result does not require behavioural arguments making sense out of it. We have identified some confusion on this point in empirical applications (e.g., Kassouf and Hoffmann, 2006).

References

Blasch, J. E., Filippini, M., and Kumar, N. (2017). Boundedly rational consumers, energy and investment literacy, and the display of information on household appliances. *Resource and Energy Economics*, In Press.

Greene, W. H. (1998). Gender Economics Courses in Liberal Arts Colleges: Further Results. *The Journal of Economic Education*, 29(4):291–300.

Greene, W. H. (2018). *Econometric analysis*. New York: Pearson, 8th edition.

Humphreys, B. R., McLeod, L., and Ruseski, J. E. (2014). Physical activity and health outcomes: evidence from canada. *Health Economics*, 23(1):33–54.

Kassouf, A. L. and Hoffmann, R. (2006). Work-Related Injuries Involving Children and Adolescents: Application of a Recursive Bivariate Probit Model. *Brazilian Review of Econometrics*, 26(1):105–126.

Maddala, G. S. (1986). *Limited-dependent and qualitative variables in econometrics*. Number 3. Cambridge university press.

Wilde, J. (2000). Identification of multiple equation probit models with endogenous dummy regressors. *Economics letters*, 69(3):309–312.

Working Papers of the Center of Economic Research at ETH Zurich

(PDF-files of the Working Papers can be downloaded at www.cer.ethz.ch/research/working-papers.html).

17/275 M. Filippini, W. Greene, N. Kumar, A. Martinez.Cruz
A note on the different interpretation of the correlation parameters in the Bivariate Probit and the Recursive Bivariate Probit

17/274 D. Basin, H. Gersbach, A. Mamageishvili, L. Schmid and O. Tejada
Election Security and Economics: It's all about Eve

17/273 J. Abrell, M. Kosch and S. Rausch
The Economic Cost of Carbon Abatement with Renewable Energy Policies

17/272 H. Gersbach and O. Tejada
Semi-Flexible Majority Rules for Public Good Provision

17/271 D. Cerruti, A. Alberini, J. Linn
Charging Drivers by the Pound: The Effects of the UK Vehicle Tax System

17/270 H. Gersbach, P. Muller, O. Tejada
A Dynamic Model of Electoral Competition with Costly Policy Changes

17/269 J. Blasch, N. Boogen, M. Filippini, N. Kumar
The role of energy and investment literacy for residential electricity demand and end-use efficiency

17/268 H. Gersbach, M.-C. Riekhof
Technology Treaties and Climate Change

17/267 Christos Karydas
The inter-temporal dimension to knowledge spillovers: any non-environmental reason to support clean innovation?

17/266 Christos Karydas, Lin Zhang
Green tax reform, endogenous innovation and the growth dividend

17/265 Daniel Harenberg, Stefano Marelli, Bruno Sudret, Viktor Winschel
Uncertainty Quantification and Global Sensitivity Analysis for Economic Models

16/264 Marie-Catherine Riekhof
The Insurance Premium in the Interest Rates of Interlinked Loans in a Small-scale Fishery

16/263 Julie Ing
Adverse selection, commitment and exhaustible resource taxation

16/262 Jan Abrell, Sebastian Rausch, and Giacomo A. Schwarz
Social Equity Concerns and Differentiated Environmental Taxes

16/261 D. Ilic, J.C. Mollet
Voluntary Corporate Climate Initiatives and Regulatory Loom: Batten Down the Hatches

16/260 L. Bretschger
Is the Environment Compatible with Growth? Adopting an Integrated Framework

16/259 V. Grossmann, A. Schaefer, T. Steger, and B. Fuchs
Reversal of Migration Flows: A Fresh Look at the German Reunification

16/258 V. Britz, H. Gersbach, and H. Haller
Deposit Insurance in General Equilibrium

16/257 A. Alberini, M. Bareit, M. Filippini, and A. Martinez-Cruz
The Impact of Emissions-Based Taxes on the Retirement of Used and Inefficient Vehicles: The Case of Switzerland

16/256 H. Gersbach
Co-voting Democracy

16/255 H. Gersbach and O. Tejada
A Reform Dilemma in Polarized Democracies

16/254 M.-C. Riekhof and J. Broecker
Does the Adverse Announcement Effect of Climate Policy Matter? - A Dynamic General Equilibrium Analysis

16/253 A. Martinez-Cruz
Handling excess zeros in count models for recreation demand analysis without apology

16/252 M.-C. Riekhof and F. Noack
Informal Credit Markets, Common-pool Resources and Education

16/251 M. Filippini, T. Geissmann, and W. Greene
Persistent and Transient Cost Efficiency - An Application to the Swiss Hydropower Sector

16/250 L. Bretschger and A. Schaefer
Dirty history versus clean expectations: Can energy policies provide momentum for growth?

16/249 J. Blasch, M. Filippini, and N. Kumar
Boundedly rational consumers, energy and investment literacy, and the display of information on household appliances

16/248 V. Britz
Destroying Surplus and Buying Time in Unanimity Bargaining

16/247 N. Boogen, S. Datta, and M. Filippini
Demand-side management by electric utilities in Switzerland: Analyzing its impact on residential electricity demand

16/246 L. Bretschger
Equity and the Convergence of Nationally Determined Climate Policies

16/245 A. Alberini and M. Bareit
The Effect of Registration Taxes on New Car Sales and Emissions: Evidence from Switzerland

16/244 J. Daubanes and J. C. Rochet
The Rise of NGO Activism

16/243 J. Abrell, Sebastian Rausch, and H. Yonezawa
Higher Price, Lower Costs? Minimum Prices in the EU Emissions Trading Scheme

16/242 M. Glachant, J. Ing, and J.P. Nicolai
The incentives to North-South transfer of climate-mitigation technologies with trade in polluting goods

16/241 A. Schaefer
Survival to Adulthood and the Growth Drag of Pollution

16/240 K. Prettner and A. Schaefer
Higher education and the fall and rise of inequality

16/239 L. Bretschger and S. Valente
Productivity Gaps and Tax Policies Under Asymmetric Trade

16/238 J. Abrell and H. Weigt
Combining Energy Networks

16/237 J. Abrell and H. Weigt
Investments in a Combined Energy Network Model: Substitution between Natural Gas and Electricity?

16/236 R. van Nieuwkoop, K. Axhausen and T. Rutherford
A traffic equilibrium model with paid-parking search

16/235 E. Balistreri, D. Kaffine, and H. Yonezawa
Optimal environmental border adjustments under the General Agreement on Tariffs and Trade

16/234 C. Boehringer, N. Rivers, H. Yonezawa
Vertical fiscal externalities and the environment