Bivariate Probit Estimation for Panel Data: a two-step Gauss-Hermite Quadrature Approach with an application to product and process innovations for France

Benoît MULKAY
Université Montpellier 1*

March 2014

Preliminary, Do not quote !

Abstract

This paper describes two methods for computing a bivariate probit model on panel data with correlated random effects. A first approach using simulated maximum likelihood has been already presented in the literature. An alternative method based on a two-step Gauss-Hermite quadrature in order to evaluate the likelihood function is proposed in this article. A simulation shows the importance to estimate the correlation in random effects and the correlation between both equations. Finally an application is performed to estimate the determinants of product or process innovations on a panel of French firms. It shows a very large correlation between individual effects.

JEL code : C33, C35, O31.

*Address : Faculté d'Economie, Avenue Raymond Dugrand CS 79606, 34960 Montpellier Cedex 2. E-mail : benoit.mulkay@univ-montp1.fr

1 Introduction

The estimation of a probit model on panel data is now usual. Many softwares propose such method of estimation which relies on individual random effects because the fixed effects approach is not valid due to the incidental parameters problem in the non-linear panel data model. In a seminal paper, Butler and Moffitt (1982) suggested to integrating the density conditional over the distribution of the individual effects in order to eliminate them by taking an average density. They proposed to use a Gauss-Hermite Quadrature to compute this integral for each individual in the panel. On the other hand, many empirical problems imply two binary variables. The classic bivariate probit model is now common for cross-section data, but no usual procedure is available for panel data where there is an individual random effect. In fact, it is interesting to estimate the correlation between the individual effects in the two equations, because it shows how the unoberved heterogeneity of individuals is correlated accross equation, while there is still a correlation between the idiosyncratic error terms in the two equations.

On the line of multivariate probit model, Lee and Oguzoglu (2007) and Kano (2008) have proposed a simulated maximum likelihood approach where the individuals effects are integrated out by computing the double integral by simulation. But this procedure could be very time-consuming even with fast modern computer. In this article, an alternative approach based on a two-step Gauss-Hermite Quadrature is used in order to compute this double integral, which should be rather in the line of the Butler and Moffitt (1982) approach. Such a method has been already investigated in the context of a Heckman selection model on panel data by by Raymond et al (2007, 2010). I adapt their method in the case of a bivariate panel data model in the section 2.

A simulation analysis is done in Section 3 in order to show the importance of taking account individual effects in estimation of a probit model on panel data. The separated estimation of the two probit models shows clearly that they are consistent due to the fact that the model is correctly specified and that the correlations between the individual effects or between the error terms are only of second order. In fact like in a seemingly unrelated regression equations model, there is only a gain in efficiency of taking account of the covariance structure of the error terms composed of an individual effect and a idiosyncratic error. However the estimation of the correlations is of interest in order to assess the effects of unobserved heterogeneity on each equation.

Finally in Section 4, we present an application of this procedure in the case of the estimation of the determinants of product and process innovations on a panel of French firms during the period 2000 - 2007. The French data are annual, but they indicate only whether a firm, with R&D expenditures, introduces a product or a process innovation during the given year. The model explaining the product or process innovations is simple because it depends only on the size of the firm and on the R&D intensity. There is a positive effect of the size of the firm of the same magnitude for both types of innovation, while there is also a positive effect of the R&D intensity, but this effect is non-linear because it is decreasing up to a R&D intensity which is roughly 50 % of the total turnover of the firm. Finally the unobserved heterogeneity affects both equation with a very high positive correlation of 90 %, even though the estimated correlation between the idiosyncratic errors terms is about 50 %. In consequence the individual characteristics have the same impact on both type of innovations.

2 The random effect bivariate probit

2.1 The bivariate probit model

Here we present briefly the bivariate probit model[1]. Thismodel is composed by 2 latent variables y_1^* and y_2^* which are explained by exogenous variables x_1 and x_2 and by possibly correlated error terms ε_1 and ε_2, normally distributed with unit variances[2]. and correlation coefficient τ :

$$\left\{ \begin{array}{l} y_1^* = x_1'\beta_1 + \varepsilon_1 \\ y_2^* = x_2'\beta_2 + \varepsilon_2 \end{array} \right. \qquad \text{where } \varepsilon = \left(\begin{array}{c} \varepsilon_1 \\ \varepsilon_2 \end{array} \right) \approx i.i.d.N\left[\left(\begin{array}{c} 0 \\ 0 \end{array} \right); \left(\begin{array}{cc} 1 & \tau \\ \tau & 1 \end{array} \right) \right]$$

If the data are observed on several individuals only, we obtain the classical bivariate probit model when the observed variables y_1 and y_2 are defined as :

$$\left\{ \begin{array}{l} y_1 = 1\,(y_1^* > 0) \\ y_2 = 1\,(y_2^* > 0) \end{array} \right.$$

where $1\,(...)$ is the indicator function with value one if the expression in parenthesis is true, and zero otherwise. The maximum likelihood estimator is then simply obtained with the classical transformation :

$$q_j = 2y_j - 1$$

such that the probability of a given choice between the 4 possible configurations of choice is :

$$\Pr\left(Y_1 = y_1, Y_2 = y_2 |\, x_1, x_2; \beta_1, \beta_2, \tau\right) = \Phi_2\left[q_1\left(x_1'\beta_1\right), q_2\left(x_2'\beta_2\right), q_1 q_2 \tau\right]$$

with $\Phi_2(\bullet)$ is the cumulative density fonction of the bivariate standard normal distribution :

$$\begin{aligned} \Phi_2\left[u_1, u_2; \tau\right] &= \int_{-\infty}^{u_1}\int_{-\infty}^{u_2} \phi_2\left(z_1, z_2; \tau\right) dz_1 dz_2 \\ &= \int_{-\infty}^{u_1}\int_{-\infty}^{u_2} \frac{1}{2\pi}\frac{1}{\sqrt{1-\tau^2}} \exp\left\{-\frac{z_1^2 + z_2^2 - 2\tau z_1 z_2}{2\left(1-\tau^2\right)}\right\} dz_1 dz_2 \end{aligned}$$

[1] See for example : Greene (2008, Section XXI.6).

[2] The classical normalization of variances to unity is done here because only the signs of the latent variables are observed. Therefore the scale does not matter.

where $\phi_2(\bullet)$ is its probability density function of a bivariate standard normal variable with correlation τ. As the N observations of the sample are independent, the log-likelihood function is given by :

$$\ln \pounds = \sum_{i=1}^{n} \ln \Phi_2 \left[q_{1,i} x'_{1,i}\beta_1, q_{2,i} x'_{2,i}\beta_2; q_{1,i} q_{2,i}\tau\right]$$

which should be maximized to obtain the maximum likelihood estimator of the bivariate probit model[3]. Greene (2008) gives the analytic first and second order conditions of the estimation problem.

When the observations come from a panel of individuals observed during a given time period (supposed here for simplicity to be the same for all individuals such that the panel is balanced), there is often an individual effect to take account of the unobserved heterogenity of the individuals. However with a non-linear model, like the probit model, if the individual effects are treated as fixed or correlated with the explanatory variables, there is an incidental parameters problem (Neyman and Scott, 1948; Lancaster, 2000; or Cameron and Trivedi; 2006). Thus we need to assume that individual effects are not correlated with the explanatory variables, and we use a random effects model with a specified distribution. These random effects are then eliminated by integrating over the distribution.

The univariate probit case has been first studied by Butler and Moffitt (1982) and Skrondal and Rabe-Hasketh (2004). We generalize this univariate random effect probit model to the case of two latent variables for $i = 1, ..., N$ individuals and $t = 1, ..., T$ time periods :

$$\left\{ \begin{array}{l} y^*_{1,it} = x'_{1,it}\beta_1 + \alpha_{1,i} + \varepsilon_{1,it} \\ y^*_{2,it} = x'_{2,it}\beta_2 + \alpha_{2,i} + \varepsilon_{2,it} \end{array} \right. \qquad \text{for } i = 1, ..., N \text{ and } t = 1, ..., T.$$

$$\text{where : } \left\{ \begin{array}{l} \varepsilon_{it} = \begin{pmatrix} \varepsilon_{1,it} \\ \varepsilon_{2,it} \end{pmatrix} \approx i.i.d.N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix} \right] \\ \alpha_i = \begin{pmatrix} \alpha_{1,i} \\ \alpha_{2,i} \end{pmatrix} \approx i.i.d.N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right] \end{array} \right.$$

Here we assume implicitly that the observations are independant over time and accross indivisuals. The explanatory variables are exogenous with respect to the error terms and with the individual random effects. This last hypothesis could be relaxed by introducing the average values of the regressors along the lines proposed by Mundlak (1978) if the individual effect can be decomposed on a linear combination of the averaged regressors plus an uncorrelated effects. The observed model is:

$$\left\{ \begin{array}{l} y_{1,it} = 1\left(y^*_{1,it} > 0\right) \\ y_{2,it} = 1\left(y^*_{2,it} > 0\right) \end{array} \right.$$

[3] See for example the estimation procedure `biprobit` in **Stata** (Hardin, 1996).

Let us define the classical transformation of the observed variables :

$$\begin{cases} q_{1,it} = 2y_{1,it} - 1 \\ q_{2,it} = 2y_{2,it} - 1 \end{cases}$$

2.2 The individual joint density function

Because of the independance of observations over time, the conditional joint density for the T observations of the i^{th} individual is:

$$f_i\left(y_i | X_i, \alpha_i, \beta, \tau\right) = \prod_{t=1}^{T} f_{it}\left(y_{it} | X_{it}, \alpha_i, \beta, \tau\right)$$

As we have assumed a normal distribution for the error terms in the latent model, the density for an observation is given as in the bivariate probit model above:

$$\begin{aligned} & \Pr\left(Y_1 = y_1, Y_2 = y_2 | x_1, x_2; \alpha_i, \beta, \tau\right) \\ = \; & \Phi_2\left(q_1\left(x_1'\beta_1 + \alpha_1\right), q_2\left(x_2'\beta_2 + \alpha_2\right); q_1 q_2 \tau\right) \end{aligned}$$

where the individual random effects are added up to the conventional observable parts of the latent functions. The joint density for an individual, conditional to the vector of the individual random effects $\alpha_i = (\alpha_{1,i}, \alpha_{2,i})$, is then:

$$f_i\left(y_i | X_i, \alpha_i, \beta, \tau\right) = \prod_{t=1}^{T} \Phi_2\left(q_{1,it}\left(x_{1,it}'\beta_1 + \alpha_{1,i}\right), q_{2,it}\left(x_{2,it}'\beta_2 + \alpha_{2,i}\right); q_{1,it} q_{2,it} \tau\right)$$

Assuming a normal disribution for these individual random effects with variances σ_1^2 and σ_2^2 respectively and a correlation coefficient ρ, the density function for the individual effects is given by :

$$\begin{aligned} g_i\left(\alpha_i | \sigma_1^2, \sigma_2^2, \rho\right) = \; & \frac{1}{2\pi} \frac{1}{\sqrt{\sigma_1^2 \sigma_2^2 \left(1 - \rho^2\right)}} \times \\ & \exp\left\{\frac{-1}{2\left(1-\rho^2\right)}\left[\left(\frac{\alpha_{1,i}}{\sigma_1}\right)^2 - 2\rho\left(\frac{\alpha_{1,i}}{\sigma_1}\right)\left(\frac{\alpha_{2,i}}{\sigma_2}\right) + \left(\frac{\alpha_{2,i}}{\sigma_2}\right)^2\right]\right\} \end{aligned}$$

This density function does not depend on observables but on the three parameters which should be estimated. The unconditional (to the individual random effects) joint density for the i^{th} individual is obtained by averaging over the distribution of these individual effects :

$$\begin{aligned}
& \ell_i\left(y_i \mid X_i, \beta, \tau, \sigma_1^2, \sigma_2^2, \rho\right) \\
= & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_i\left(y_i \mid X_i, \alpha_i, \beta, \tau\right) \times g_i\left(\alpha_i \mid \sigma_1^2, \sigma_2^2, \rho\right) d\alpha_{1,i} d\alpha_{2,i} \\
= & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty}\left[\prod_{t=1}^T \Phi_2\left(q_{1,it}\left(x_{1,it}^{\prime} \beta_1+\alpha_{1,i}\right), q_{2,it}\left(x_{2,it}^{\prime} \beta_2+\alpha_{2,i}\right) ; q_{1,it} q_{2,it} \tau\right)\right] \times \\
& \frac{1}{2\pi} \frac{1}{\sqrt{\sigma_1^2 \sigma_2^2\left(1-\rho^2\right)}} \times \qquad (1) \\
& \exp\left\{\frac{-1}{2\left(1-\rho^2\right)}\left[\left(\frac{\alpha_{1,i}}{\sigma_1}\right)^2-2\rho\left(\frac{\alpha_{1,i}}{\sigma_1}\right)\left(\frac{\alpha_{2,i}}{\sigma_2}\right)+\left(\frac{\alpha_{2,i}}{\sigma_2}\right)^2\right]\right\} d\alpha_{1,i} d\alpha_{2,i}
\end{aligned}$$

2.3 Decomposition of the double integral

The evaluation of the individual likelihood function (1) requires the computation of a double integral. Lee and Oguzoglu (2007) and Kano (2008) have proposed a method of computation by simulation where $\alpha_{1,i}$ and $\alpha_{2,i}$ are randomly drawn in the bivariate normal distribution[4]. The individual joint density (unconditional to the individual random effects is approximated by :

$$\begin{aligned}
& \ell_i\left(y_i \mid X_i, \beta, \tau, \sigma_1^2, \sigma_2^2, \rho\right) \\
= & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_i\left(y_i \mid X_i, \alpha_i, \beta, \tau\right) \times g\left(\alpha_i \mid \sigma_1^2, \sigma_2^2, \rho\right) d\alpha_{1,i} d\alpha_{2,i} \\
\simeq & \frac{1}{R} \sum_{r=1}^R\left[\prod_{t=1}^T \Phi_2\left(q_{1,it}\left(x_{1,it}^{\prime} \beta_1+a_{1,i}^{(r)}\right), q_{2,it}\left(x_{2,it}^{\prime} \beta_2+a_{2,i}^{(r)}\right), q_{1,it} q_{2,it} \tau\right)\right] \qquad (2)
\end{aligned}$$

where $\left(a_{1,i}^{(r)}\right)$ and $\left(a_{2,i}^{(r)}\right)$ are R random draws in a bivariate normal distribution:

$$\left(\begin{array}{c} a_{1,i}^{(r)} \\ a_{1,i}^{(r)} \end{array}\right) \sim i.i.d.N\left[\left(\begin{array}{c} 0 \\ 0 \end{array}\right);\left(\begin{array}{cc} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_1^2 \end{array}\right)\right]$$

However the computation should be very time-consuming and imprecise even though we use modern simulator like GHK or Halton simulators, because we need to compute R cumulative density function with a large value of R in order to obtain sufficient precision in the log-likelihood function.

Instead we use the two-step Gauss-Hermite quadrature technique originally proposed in a couple of papers by Raymond *et al.* (2007, 2010) in the case of an

[4] Miranda (2010) suggests the same procedure in an unpublished paper presented at the Mexican Stata Conference in 2010.

Heckman sample selection model on panel data. This method relies on a decomposition of the two-dimensional normal distribution for the individual effects into a one-dimensional marginal distribution and a one-dimensional conditional distribution.

The unconditional joint density for the i^{th} individual is rewritten as:

$$\begin{aligned}
&\ell_i\left(y_i|\, X_i, \beta, \tau, \sigma_1^2, \sigma_2^2, \rho\right) \\
=& \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_i\left(y_i|\, X_i, \alpha_i, \beta, \tau\right) \times \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)}} \times \\
& \exp\left\{\frac{-1}{2\left(1-\rho^2\right)}\left[\left(\frac{\alpha_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{\alpha_1}{\sigma_1}\right)\left(\frac{\alpha_2}{\sigma_2}\right) + \left(\frac{\alpha_2}{\sigma_2}\right)^2\right]\right\} d\alpha_1 d\alpha_2 \\
=& \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_i\left(y_i|\, X_i, \alpha_i, \beta, \tau\right) \times \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)}} \times \\
& \exp\left\{\frac{-1}{2\left(1-\rho^2\right)}\left[\left(\frac{\alpha_{1,i}}{\sigma_1}\right)^2 - 2\rho\left(\frac{\alpha_{1,i}}{\sigma_1}\right)\left(\frac{\alpha_{2,i}}{\sigma_2}\right)\right]\right\} \times \exp\left[-\frac{1}{2}\frac{\left(\alpha_{2,i}/\,\sigma_2\right)^2}{1-\rho^2}\right] d\alpha_{1,i} d\alpha_{2,i}
\end{aligned}$$

which can be in turn rewritten as:

$$\ell_i\left(y_i|\, X_i, \beta, \sigma_1^2, \sigma_2^2, \rho\right) = \int_{-\infty}^{+\infty} H_i\left(\alpha_{2,i}\right) \times \exp\left[-\frac{1}{2}\frac{\left(\alpha_{2,i}/\,\sigma_2\right)^2}{1-\rho^2}\right] d\alpha_{2,i} \qquad (3)$$

with

$$\begin{aligned}
H_i\left(\alpha_{2,i}\right) =& \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)}}\int_{-\infty}^{+\infty} f_i\left(y_i|\, X_i, \alpha_i, \beta, \tau\right) \times \\
& \exp\left\{\frac{-1}{2\left(1-\rho^2\right)}\left[\left(\frac{\alpha_{1,i}}{\sigma_1}\right)^2 - 2\rho\left(\frac{\alpha_{1,i}}{\sigma_1}\right)\left(\frac{\alpha_{2,i}}{\sigma_2}\right)\right]\right\} d\alpha_{1,i}
\end{aligned}$$

Let us evaluate this last function by using a gauss-Hermite Quadrature by doing a change in variable such that $\left(\alpha_1/\,\sigma_1\right) = z_1\sqrt{2\left(1-\rho^2\right)}$ with $d\alpha_1 = \sigma_1\sqrt{2\left(1-\rho^2\right)}dz_1$ such that[5]:

$$\begin{aligned}
H\left(\alpha_2\right) =& \frac{1}{2\pi}\frac{\sigma_1\sqrt{2\left(1-\rho^2\right)}}{\sqrt{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)}}\int_{-\infty}^{+\infty} \ell_i\left(y_i|\, X_i, z_1\sigma_1\sqrt{2\left(1-\rho^2\right)}, \alpha_2, \beta, \tau\right) \times \\
& \exp\left\{-\frac{1}{2}\frac{z_1^2 2\left(1-\rho^2\right)}{1-\rho^2}\right\} \times \exp\left\{\frac{\rho}{\left(1-\rho^2\right)} z_1\sqrt{2\left(1-\rho^2\right)}\left(\frac{\alpha_2}{\sigma_2}\right)\right\} dz_1 \\
=& \frac{1}{\pi\sqrt{2\sigma_2^2}}\int_{-\infty}^{+\infty} f_i\left(y_i|\, X_i, z_1\sigma_1\sqrt{2\left(1-\rho^2\right)}, \alpha_2, \beta, \tau\right) \times \\
& \exp\left\{\frac{\rho\sqrt{2}}{\sqrt{1-\rho^2}}\left(\frac{\alpha_2}{\sigma_2}\right) z_1\right\} \times \exp\left\{-z_1^2\right\} dz_1.
\end{aligned}$$

[5]We drop the individual index i for the clarity of the exposition.

This is a Gaussian integral which can be approximated by a Gauss-Hermite quadrature with weights ω_m and abscissas a_m for M integration points ($m = 1, ..., M$)[6] :

$$\int_{-\infty}^{+\infty} f(z)\, e^{-z^2} dz \simeq \sum_{m=1}^{M} \omega_m f(a_m)$$

Thus the $H(\alpha_2)$ function is approximated by :

$$H_i(\alpha_{2,i}) \simeq \frac{1}{\pi\sqrt{2\sigma_2^2}} \sum_{m=1}^{M} \omega_m f_i \left(y_i | X_i, a_m \sigma_1 \sqrt{2(1-\rho^2)}, \alpha_{2,i}, \beta, \tau \right) \exp \left[\frac{\rho\sqrt{2}}{\sqrt{1-\rho^2}} \left(\frac{\alpha_{2,i}}{\sigma_2} \right) a_m \right].$$

Now the second step of the procedure is to introduce this function in the individual joint density $\ell_i \left(y_i | X_i, \beta, \tau, \sigma_1^2, \sigma_2^2, \rho \right)$ above (3) with a second change in variables $(\alpha_2/\sigma_2) = z_2\sqrt{2(1-\rho^2)}$ with $d\alpha_2 = \sigma_2\sqrt{2(1-\rho^2)} dz_2$ to obtain:

$$\begin{aligned}
& \ell_i \left(y_i | X_i, \beta, \tau, \sigma_1^2, \sigma_2^2, \rho \right) \\
= & \int_{-\infty}^{+\infty} H(\alpha_2) \exp \left[-\frac{1}{2} \frac{(\alpha_2/\sigma_2)^2}{1-\rho^2} \right] d\alpha_{2,i} \\
= & \frac{1}{\pi\sqrt{2\sigma_2^2}} \int_{-\infty}^{+\infty} \sum_{m=1}^{M} \omega_m f_i \left(y_i | X_i, a_m \sigma_1 \sqrt{2(1-\rho^2)}, \alpha_2, \beta, \tau \right) \\
& \times \exp \left[\frac{\rho\sqrt{2}}{\sqrt{1-\rho^2}} \left(\frac{\alpha_2}{\sigma_2} \right) a_m \right] \times \exp \left[-\frac{1}{2} \frac{(\alpha_2/\sigma_2)^2}{1-\rho^2} \right] d\alpha_2 \\
= & \frac{\sigma_2\sqrt{2(1-\rho^2)}}{\pi\sqrt{2\sigma_2^2}} \int_{-\infty}^{+\infty} \sum_{m=1}^{M} \omega_m f_i \left(y_i | X_i, a_m \sigma_1 \sqrt{2(1-\rho^2)}, z_2 \sigma_2 \sqrt{2(1-\rho^2)}, \beta, \tau \right) \\
& \times \exp \left[\frac{\rho\sqrt{2}}{\sqrt{1-\rho^2}} z_2 \sqrt{2(1-\rho^2)} a_m \right] \times \exp \left[-\frac{1}{2} \frac{z_2^2 2(1-\rho^2)}{1-\rho^2} \right] dz_2 \\
= & \frac{\sqrt{1-\rho^2}}{\pi} \int_{-\infty}^{+\infty} \sum_{m=1}^{M} \omega_m \ell_i \left(y_i | X_i, a_m \sigma_1 \sqrt{2(1-\rho^2)}, z_2 \sigma_2 \sqrt{2(1-\rho^2)}, \beta, \tau \right) \\
& \times \exp[2\rho a_m z_2] \times \exp\left[-z_2^2\right] dz_2
\end{aligned}$$

A second Gauss-Hermite quadrature can be used to compute this Gaussian integral. For P integration points ($p = 1, ..., P$), we have the weights ω_p and the abscissas a_p. Finally the individual joint density unconditional to the individual effects can be approximated by :

[6] The more the number of points, the more precise is the approximation. Genrally the number of points is set to $8, 12$ or 16 (see Cameron and Trivedi, 2005, Section XII.3.1). The values of weights ω_m and abscissas a_m can be found in mathematical textbooks.

$$
\begin{aligned}
& \ell_i\left(y_i | X_i, \beta, \tau, \sigma_1^2, \sigma_2^2, \rho\right) \\
\simeq\ & \frac{\sqrt{(1-\rho^2)}}{\pi} \sum_{p=1}^{P} \sum_{m=1}^{M} \omega_p \omega_m \exp\left[2\rho a_m a_p\right] \left(\prod_{t=1}^{T} \Phi_2\left(q_1 u_{1,m}; q_2 u_{2,p}; q_1 q_2 \tau\right) \right)
\end{aligned} \tag{4}
$$

where the arguments of the bivariate cumulative density function are :

$$
\begin{aligned}
u_{1,m} &= x_1' \beta_1 + a_m \sigma_1 \sqrt{2\left(1-\rho^2\right)} \\
u_{2,p} &= x_2' \beta_2 + a_p \sigma_2 \sqrt{2\left(1-\rho^2\right)}
\end{aligned}
$$

Finally as the individuals are independent, the log-likelihood function should be expressed as:

$$
\begin{aligned}
\ln \pounds &= \sum_{i=1}^{N} \ln \ell_i\left(y_i | X_i, \beta, \tau, \sigma_1^2, \sigma_2^2, \rho\right) \\
&= -N \ln(\pi) + \frac{N}{2} \ln\left(1-\rho^2\right) + \\
& \sum_{i=1}^{N} \ln \left(\sum_{p=1}^{P} \sum_{m=1}^{M} \omega_p \omega_m \exp\left[2\rho a_m a_p\right] \prod_{t=1}^{T} \Phi_2\left(q_{1,i} u_{1,m,i}; q_{2,i} u_{2,m,i}; q_{1,i} q_{2,i} \tau\right) \right)
\end{aligned} \tag{5}
$$

In order to maximize this log-likelihood function, we can use the usual transformations for the correlation coefficients:

$$
\begin{cases}
\rho^* = a\tanh\rho = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right) \\
\tau^* = a\tanh\tau = \frac{1}{2}\ln\left(\frac{1+\tau}{1-\tau}\right)
\end{cases}
$$

or

$$
\begin{cases}
\rho = \frac{\exp(2\rho^*)-1}{\exp(2\rho^*)+1} \\
\tau^* = \frac{\exp(2\tau^*)-1}{\exp(2\tau^*)+1}
\end{cases}
$$

At each evaluation of the likelihood function, it is necessary to compute $N \times M \times P$ cumulative density functions of the bivariate normal variables Φ_2 with this two-step quadrature, which seems much more reasonable relative to the computation of $N \times R^2$ cumulative density functions for the simulated method. In fact we should have a sufficiently good approximation with $M = P = 12$ points in the Gauss-Hermite quadrature, even though we shoud take at least $R = 200$ points for the computation by simulation with less precision. The two procedures of estimation of the bivariate probit model by maximum likelihood have been written in a Stata program either with the simulated maximum likelihood or with the Gauss-Hermite quadrature[7].

[7] These programs uses the maximum likelihood procedures in Stata by Gould *et al.* (2010).

3 A simulation

A simulation of the procedures for the estimation of the bivariate probit model has been performed in order to assess the effect of neglecting the correlation between the two equations, and between the unobserved heterogeneity in each equation. A set of observations for N individuals during T periods has been generated for a bivariate latent process:

$$\begin{cases} y_1^* = \beta_{1,0} + \beta_{1,1}x_1 + \beta_{1,2}x_2 + \alpha_1 + \varepsilon_1 \\ y_2^* = \beta_{2,0} + \beta_{2,1}x_1 + \beta_{2,2}x_2 + \alpha_2 + \varepsilon_2 \end{cases}$$

$$\text{where } \varepsilon \quad = \quad \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \approx i.i.d.N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix} \right]$$

where the exogenous variables x_1 and x_2 have been drawn independently for each observations in a standard normal distribution. The individual effects α_1 and α_2 have been also drawn into a bivariate normal distribution with correlation ρ:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \approx i.i.d.N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right].$$

Then the observable dependent variables are constructed on the basis of the sign of the corresponding latent variables:

$$\begin{cases} y_1 = 1\,(y_1^* > 0) \\ y_2 = 1\,(y_2^* > 0) \end{cases}$$

In the following simulations, the number of individuals has been set to 1 000 with 10 periods for each individuals, such that there are 10 000 observations in the panel data set which corresponds to the usual size of such data. The true structural parameters in the model are the following : $\beta_{1,0} = 0.50, \beta_{1,1} = 1.00, \beta_{1,2} = 0.00$ and $\beta_{2,0} = -0.50, \beta_{2,1} = -0.50, \beta_{2,2} = 1.00$. Therefore the second explanatory variable appears only in the second equation. The correlation coefficient of the error terms has been set to $\tau = 0.50$, the same value has the correlation coefficient between the individual random effects : $\rho = 0.50$, while the standard deviation of these individual effects are the same: $\sigma_1 = \sigma_2 = 2.00$.

The observed patterns of response in this simulated model is the shown in the Table 1.This simulated data set exhibits an association between both dependent variable with a Kendall's-t_b measure of association of 0.222 with a standard error of 0.010, as well a Pearson Chi-squared of 491.55 showing clearly a positive significant association between the two observed dependent variables. Moreover the tetrachoric correlation is 0.349 with a standard error 0.015 which is less than the assumed correlation between the error terms in the latent model.

		y_2		
		0	1	Total
y_1	0	29.7 %	12.4 %	42.0 %
	1	28.1 %	29.9 %	58.0%
	Total	57.7 %	42.3 %	100 %

Table 1 : Contingency table of the binary variables in the simualted model.

This simulated data set exhibits an association between both dependent variable with a Kendall's-t_b measure of association of 0.222 with a standard error of 0.010, as well a Pearson Chi-squared of 491.55 showing clearly a positive significant association between the two observed dependent variables. Moreover the tetrachoric correlation is 0.349 with a standard error 0.015 which is less than the assumed correlation between the error terms in the latent model.

The model is estimated by a pooled bivariate probit method where there are no individual effects as a benchmark for estimations. Then it is estimated using the Gauss-Hermite Quadrature (with 12 points) allowing individual effects. We proceed to four estimations : the first one (estimation 1) with the individual random effects but with a zero correlation between error terms ($\tau = 0$) and a zero correlation between the individual effects ($\rho = 0$), the second estimation (2) allows for an estimated correlation between the error terms (τ), wile the third estimation (3) allows only a correlation between the individual effects (ρ). Finally the last estimation (4) is the complete model where both correlations must be estimated. The standard likelihood ratio tests are performed in order to verify the assumption about the individual effects and the correlations in the model.

The Gauss-Hermite Quadrature procedure with 12 integration points is here faster by 40 % than the simulated maximum likelihhod procedures performed on the same dataset and on the same computer. Even though the convergence is quite fast in three or four iterations starting with the initial values from the two univariate panel probit estimations, it takes hovever between 7 minutes (for the first estimation) to 14 minutes (for the last estimation) to perform such a regression[8] on 10 000 observations for a model with only 3 parameters in each equation !

The benchmark estimation is clearly biased for the structural parameters of each equation because there is no individual effects. Only a correlation between the idiosyncratic error terms is estimated with an estimated value (0.511) close to the theoretical correlation (0.50). Let us remark that the parameter estimates are less than the half of their theoretical values. A likelihood ratio test rejects clearly this hypothesis of no individual random effects. Introducing individual random effects in the estimation but with no correlation is equivalent to two distinct estimation of a random effect probit model for each equation. The structural parameter estimates are now close to their theoretical value, taking

[8] The estimation are performed on a Dell OptiPlex 9010 with a i7 Intel processor running at 3.4 Ghz. The procedures are written in a standard code for maximum likelihhod estimation with Stata 12 software.

account their standard errors. This is rather the case for the second equation, while the first one presents estimates a litle bit smaller than theitr theoretical values. However the estimated standard deviations of the individual effects are lower than expected for both equations. The likelihood ratio tests of the correlations between individual effects and/or between the error terms in the model clearly accept the presence of such correlations in the estimations. Moreover these estimated correlations have a very small estimated standard error, even though they are non-linear transformations of the estimated parameters in constructing interval confidence foor these correlations.

	Benchmark	**(1)**	**(2)**	**(3)**	**(4)**
			Equation 1		
$\beta_{1,0}$	0.212	0.429	0.440	0.436	0.446
[= 0.50]	*(0.013)*	*(0.047)*	*(0.047)*	*(0.052)*	*(0.051)*
$\beta_{1,1}$	0.408	0.928	0.927	0.942	0.939
[= 1.00]	*(0.014)*	*(0.027)*	*(0.026)*	*(0.027)*	*(0.027)*
$\beta_{2,1}$	−0.005	−0.005	−0.004	−0.004	−0.003
[= 0.00]	*(0.013)*	*(0.020)*	*(0.020)*	*(0.020)*	*(0.020)*
			Equation 2		
$\beta_{1,0}$	−0.218	−0.523	−0.507	−0.545	−0.511
[= −0.50]	*(0.013)*	*(0.067)*	*(0.061)*	*(0.067)*	*(0.065)*
$\beta_{1,1}$	−0.220	−0.489	−0.493	−0.495	−0.497
[= −0.50]	*(0.014)*	*(0.022)*	*(0.022)*	*(0.023)*	*(0.022)*
$\beta_{2,1}$	0.465	1.016	1.008	1.029	1.022
[= 1.00]	*(0.014)*	*(0.028)*	*(0.028)*	*(0.028)*	*(0.028)*
		Standard Error of Individual Effects			
σ_1	0	1.777	1.737	2.002	1.995
[= 2.00]	–	*(0.098)*	*(0.095)*	*(0.128)*	*(0.127)*
σ_2	0	1.668	1.637	1.902	1.895
[= 2.00]	–	*(0.092)*	*(0.090)*	*(0.124)*	*(0.122)*
			Correlations		
τ	0.511	0	0.534	0	0.476
[= 0.50]	*(0.014)*	–	*(0.046)*	–	*(0.036)*
ρ	0	0	0	0.550	0.536
[= 0.50]	–	–	–	*(0.027)*	*(0.027)*
Log Likelihood	−11972.8	−7896.7	−7816.4	−7762.9	−7688.3

Standard errors of estimates in parenthesis.

True value of parameters in squared brackets in first column.

Table 2 : Simulation Results

If a correlation between the error terms in both equations is allowed ($\tau \neq 0$), the estimated results are closer from the theoretical values, while the standard deviation of the individual effect are again under-estimated. In the opposite if only a correlation between individual random effects is allowed in the estimation ($\rho \neq 0$), there are small changes in the structural parameters estimates, even though the estimated standard deviations of the individual effects are now close from their theoretical values. The same conclusions are obtained in the full model where both correlations are estimated. All estimated parameters are now very close from their theoretical values, and the hypothesis of no correlations between individual effects and between error terms is clearly rejected by the likelihood ratio tests.

4 An application to product and process innovations

In this section, we investigate the behavior of product and process innovations on a panel of French firms on the period 1999 - 2007. The data comes from the annual R&D surveys conllected each year by the Ministry of Research. The 1999 reform of the R&D surveys in France introduced two new question s about the product or the process innovations. These question are stated as :

"During the year, did your enterprise or your group introduce new or significantly improved goods coming from the R&D activity of your firm?"
(Yes or No)
"During the year, did your enterprise or your group introduce new or significantly improved methods of manufacturing or producing goods or services coming from the R&D activity of your firm?"
(Yes or No)

These questions are slightly different from the usual Community Innovation Survey (CIS) questionnaire because in the latter the time period is prolonged over 3 years. For examples in the CIS 2004 questions, the first words are replaced by "During the three years 2002 to 2004,...". Moreover in the French R&D surveys, only innovations coming from the R&D done by the firm are considered. That excludes the innovations which were introduced without any R&D effort. On the other hand, the product or process innovations can be done by another firm in the group. This is why the answers to the CIS surveys and the R&D surveys are not directly comparable. But the most important difference is that in CIS surveys, the innovations are accounted for on the three years period.

A second problem arises from the fact that firms has many difficulties to disentangle product or process innovations, even though the definitions from the Oslo manual are quite precise (see the discussion in Mairesse and Mohnen, 2001). When a firm introduces a new product on the market, it changes and improves also the methods of production. Therefore, the product and process innovations is linked at the firm level. Even though this problem of measurement is a serious one, we will consider both types of innovations in the following.

While there are some firms which innovates only in product or in process, the statistical difference between both types of innovations are thin. There are also cross relationships between product and process innovations.

The sample of the French R&D surveys covers a 8 years period : from 2000 to 2007. Only firms with at least 4 consecutive years of data are retained in the sample. There are 2 174 firms, corresponding to 12 506 observations in the unbalanced sample. 61.5 % of firms report an innovation in a new product during the year, while there are 60.3 % of firms indicating a process innovation. But large firms are more innovative than smaller firms. When the share of innovators are weighted by employment, the rate of innovation rises to 81 % for both product and process innovations. In fact about 60 % of small and medium-sized firms report an innovation, either in product or in process, while 75 % of large firms (more than 2000 employees) introduce an innovation during a given year.

		Process Innovation		
		NO	YES	TOTAL
Product	NO	26.7 %	11.8 %	38.5 %
Innovation	YES	13.0 %	48.5 %	61.5 %
	TOTAL	39.7 %	60.3 %	100 %

12 506 observations, 2 174 firms, 2000 - 2007.

Table 3 : Share of Product and Process Innovators in France

There is a positive and large association between product and process association. Nearly half of the observations in the sample lead to both types of innovation, while routhly a quarter of the sample reports no innovation at all, neither in product nor in process, even though the firms are doing R&D during the year. Finally 12 % of observations show only a product innovation, while 13 % only a process innovation. The Kendall's-τ_B measure of association is 0.479 with an asymptotic standard error of 0.007 showing a large and positive association between both types of innovations. Finally the tetrachoric correlation is 0.689 with a standard error 0.008. This clearly demonstrates the link between both type of innovations at the firm level. But this high correlation can be due to the unobserved characteristics of the firm, or rather to an idiosyncratic shock affecting both innovations at each period. We will estimate a simple bivariate probit model determining each type of innovations at the firm level to illustrate which correlations are the most important at the firm level.

In this simple model, the product or the process innovations are determined by the size of the firm, measured by the log of its total employment: $\log(L)$, and by the R&D intensity: R/Y, i.e. the total R&D expenditure divided by the total turnover of the firms. The squared value of the R&D intensity $(R/Y)^2$ is also introduced in the model in order to capture a non linear effect of the R&D intensity[9]. A full set of time dummies is also considered in the estimation. They

[9] A non-linear effect of the size of the firm have been tested with the squared log of total employment. But its parameter estimates is never significantly different from zero.

are always jointly significant in all the estimations below. Table 4 shows the main result of the analysis for the classical Probit model without any individual effect and for the Panel Probit model with individual random effects. The usual univariate estimations are done separately for each equations, while the bivariate estimations are presented with both correlations between individual effects and between the error terms. This last estimations is done by using our double Gauss-Hermite Quadrature procedures using 12 integration points for both dimensions. It takes more than 140 minutes to obtain the convergence to a maximum for the log-likelihood function but after only 5 iterations.

	PROBIT		**PANEL PROBIT**	
	UNIVARIATE	BIVARIATE	UNIVARIATE	BIVARIATE
	PRODUCT INNOVATION			
$\log(L)$	0.084	0.088	0.114	0.114
	(0.007)	*(0.007)*	*(0.015)*	*(0.014)*
(R/Y)	1.585	1.630	1.835	1.658
	(0.180)	*(0.179)*	*(0.317)*	*(0.307)*
$(R/Y)^2$	−1.550	−1.583	−1.749	−1.554
	(0.180)	*(0.180)*	*(0.310)*	*(0.300)*
Const.	−0.579	−0.596	−0.780	−0.758
	(0.059)	*(0.060)*	*(0.100)*	*(0.098)*
	PROCESS INNOVATION			
$\log(L)$	0.085	0.090	0.114	0.117
	(0.007)	*(0.007)*	*(0.015)*	*(0.015)*
(R/Y)	1.484	1.531	1.519	1.443
	(0.181)	*(0.180)*	*(0.320)*	*(0.311)*
$(R/Y)^2$	−1.560	−1.598	−1.572	−1.493
	(0.181)	*(0.181)*	*(0.313)*	*(0.303)*
Const.	−0.821	−0.862	−1.075	−1.094
	(0.060)	*(0.061)*	*(0.102)*	*(0.100)*
	STANDARD DEVIATIONS AND CORRELATIONS			
σ_1	0	0	0.990	0.508
	–	–	*(0.028)*	*(0.028)*
σ_2	0	0	1.001	0.526
	–	–	*(0.029)*	*(0.029)*
τ	0	0.714	0	0.464
	–	*(0.009)*	–	*(0.016)*
ρ	0	0	0	0.899
	–	–	–	*(0.005)*
Log.Likelihood	−16231.68	−14699.08	−14308.14	−13087.34

12 506 Observations, 2 174 Firms, 2000 - 2007

Standard errors in parenthesis. Full set of time dummies not reported here

Table 4 : Parameter Estimates

All the parameters estimates are highly significant because there is a lot of observations (12 506) in the sample, while there are up to 26 parameters to estimate in the full bivariate model. Moreover the likelihood ratio tests reject clearly the assumptions of the absence of individual effects, and the zero correlations between these individual effects or between the equations. The introduction of individual effects rises the effect of the size on the innovations in product or in process. The ffect of the R&D intensity is positive but decreasing for all the estimations because the parameter of the level is positive, even though the parameter of its squared value is negative. In consequence the effect of R&D intensity on innovations increases up to a maximum which is routhly 50 % of the total turnover of the firm. The change of the parameter estimates for the R&D intensity is mixed according to the methods of estimation. The bivariate probit model are similar for the product innovation, while its effect seems to be a lit bit smaller for the process innovations with the panel bivariate probit estimation. The estimated standard deviations of the individual effects are smaller when we take account of the correlation between these effects. They are roughly divided by 2, even though the correlation between the individual effects is very large with an estimates of $\widehat{\rho} = 0.90$. Therefore the unobserved individual characteristics of the firm seems to affect in the same way the probability to innovate in product and in process. Finally the bivariate panel probit estimation allows to disentangle the correlation between the individual effects from the correclation in the idiosyncratic error terms between the two equations, which is precisely estimated with $\widehat{\tau} = 0.46$.

The Table 5 presents the computation of the marginal effect computed at the mean value in the sample. The effect of the log employment is small but if the size of a firm is twice the size of another firm (then $\log(L)$ increases of 0.69), the probability to innovate in product or in process will be higher by 2.6 %.

	PROBIT		**PANEL PROBIT**	
	UNIVARIATE	BIVARIATE	UNIVARIATE	BIVARIATE
		PRODUCT INNOVATION		
$\log(L)$	0.029	0.030	0.037	0.037
(R/Y)	0.548	0.563	0.602	0.543
$(R/Y)^2$	-0.536	-0.547	-0.574	-0.509
		PROCESS INNOVATION		
$\log(L)$	0.029	0.031	0.037	0.038
(R/Y)	0.510	0.526	0.494	0.467
$(R/Y)^2$	-0.536	-0.549	-0.511	-0.484

12 506 Observations, 2 174 Firms, 2000 - 2007

Table 5 : Marginal Effects at the Mean

Using the bivariate panel probit estimation, the effect of the R&D intensity is also positive and reach a maximum at 54 % for the product innovation and

48 % for the process innovation. Comparing the case where the firm does not perform R&D with these maximum points, the probability of an innovation in product increases by 14.5 %, while the probability of a process innovation will be higuer by 11.3 %. These figures seems to be reasonable because about 60 % of firms in the sample innovates in product or in process.

5 Conclusions

In this article, an alternative metod of estimation of a bivariate probit model on panel data is presented. In such bivariate probit model on panel data, the likelihood function implies to integrate the density conditional over the distribution of the individual random effects in order to eliminate them by taking an average density. In the literature, some papers use simulations in order to compute the double integral. The alernative method relies on a double Gauss-Hermite quadrature procedure in order to evaluate the double integral. This paper develops the log-likelihood function in this case and a program is written in Stata to estimate such model. This program should be optimized in the future in order to reduce estimation time, may be by using an adaptative Gauss-Hermite procedure.

On panel data , it is important to introduce individial specific effects in order to avoid the omitted variable bias. This is shown in a simulation exercise where the pooled bivariate probit model is clearly rejected when there is no individual effects in estimation. The separated estimation of the two probit models is clearly consistent due to the fact that the model is correctly specified and that the correlations between the individual effects or between the error terms are only of second order. However a bivariate probit model allows also to estimate consistently the correlation between the individual random effect and between the idiosyncratic error terms in the 2 equations model. But the procedure should be long even thouh the number of iterations is reduced.

This procedure is applied in the case of the estimation of the determinants of product and process innovations on a panel of French firms during the period 2000 - 2007. Here the model explaining the product or process innovations is simple because it depends only on the size of the firm and positively on the R&D intensity. There is a positive effect of the size of the firm of the same magnitude for both types of innovation, while there is also a positive effect of the R&D intensity, but this effect is non-linear because it is decreasing up to a R&D intensity which is roughly 50 % of the total turnover of the firm. Finally the estimated correlation between the idiosyncratic errors terms is about 50 % indicating that a shock affect in the same sense with a high magnitude both types of innovations.

The unobserved heterogeneity also affects both product and process innovations with a very high positive correlation of 90 %, whicjh can be due to the fact that our model is very simple. The firm's unobserved characteristics may lead to a firm's innovative behaviour for both innovations, but these characteristics should come from the internal organization of the firm or from the market on which it operates. A further investigation of these determinants should be on the next agenda of research. The large correlated effects could be also the sign of a high persistence of innovative behaviour at the firm's level. The firm's characteristics can also affet persistenly the product or process innovations. We should investigate the persistence of this innovative behaviour in a following paper.

References

BUTLER, J. S. and Robert MOFFITT (1982) : "A computationnaly Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model", *Econometrica*, 50(3), May 1982, 761-764.

CAMERON Colin A. and Pravin K. TRIVEDI (2005) : *Microeconometrics: Methods and Applications*, Cambridge University Press.

GOULD, William W., Jeffrey S. PITBLADO, and Brian P. POI (2010) : *Maximum Likelihood Estimation with Stata (4th edition)*, College

Station, TX: Stata Press.

GREENE, William (2003) : *Econometric Analysis* (5th edition), Prentice-Hall - Pearson Education.

HARDIN, James W. (1996) : "sg61: Bivariate probit models", *Stata Technical Bulletin Reprints*, vol. 6, pp. 152–158.

KANO Shigeki (2008) : "Like Husband, Like Wife: A Bivariate Dynamic Probit Analysis of Spousal Obesities", mimeo, Osaka Prefecture University.

LANCASTER, Tony (2000) : "The Incidental Parameter Problem since 1948", *Journal of Econometrics*, 95, 391-413.

LEE, Wang-Sheng and Umut OGUZOGLU (2007) : "Well-Being and Ill-Being: A Bivariate Panel Data Analysis", IZA Discussion Paper Series No. 3108, October 2007.

MAIRESSE, Jacques and Pierre MOHNEN (2001) : "Using innovation Surveys for Econometric Analysis", NBER Working paper No. 15857.

MIRANDA, Alfonso (2010) : "Dynamic Bivariate Probit Models for Panel Data", paper presented at the Mexican Users Group Meeting, available at http://fmwww.bc.edu/repec/msug2010/mex10sug_miranda.pdf

MUNDLAK, Yairn (1978) : "On the Pooling of Time Series and Cross Section Data", *Econometrica*, 46, 69-85.

NEYMAN, J. and Elizabeth SCOTT (1948) : "Consistent Estimates Based on Partially Consistent Observations", *Econometrica*, 16(1), January 1948, 1-32.

RAYMOND Wladimir, Pierre MOHNEN, Franz PALM, and Sybrand SCHIM VAN DER LOEFF (2007) : "The Behavior of the Maximum Likelihood Estimator of Dynamic Panel Data Sample Selection Models", CESIfo Working paper No. 1992.

RAYMOND Wladimir, Pierre MOHNEN, Franz PALM, and Sybrand SCHIM VAN DER LOEFF (2010) : "Persistence of Innovation in Dutch Manufacturing: Is it Spurious ?", *Review of Economics and Statistics*, 92(3), August 2010, 495-504.

SKRONDAL, Anders and Sophia RABE-HASKETH (2004) : *Generalized Latent Variable Modelling: Multilevel, Longitudinal and Structural Equation Models*, Boca Raton, FL, Chapman and Hall.