NOTE ON THE FIML ESTIMATION OF THE THREE-EQUATION MODEL WITH MIXED POISSON AND QUALITATIVE VARIABLES

ALBERTO HOLLY

1. Notations and preliminary results

Consider a two-dimensional random vector $X = (X_1, X_2)'$ such that

$$\mathcal{D}(X) = \mathcal{N}(0, R) \tag{1.1}$$

where

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \qquad (\rho^2 < 1). \tag{1.2}$$

Let $\Psi(\rho, a, b)$ be the function defined by

$$\Psi(\rho, a, b) = \frac{1}{2\pi} \int_0^\rho \exp\left(-\frac{1}{2}\frac{a^2 + b^2 - 2tab}{1 - t^2}\right) \frac{dt}{\sqrt{1 - t^2}} \tag{1.3}$$

Hereafter we use the standard notation $\Phi(c) = \int_{-\infty}^{c} \varphi(x)dx$ and $\varphi(x) = (2\pi)^{-1/2} \exp(-x^2/2)$.

It has been shown in Lazard Holly and Holly (2002) (see also Huguenin (2004) and Huguenin, Pelgrin and Holly (2013)) that

$$P\{X_1 \geqq c_1; X_2 \geqq c_2\} = [1 - \Phi(c_1)]\,[1 - \Phi(c_2)] + \Psi(\rho, c_1, c_2) \tag{1.4}$$

and

$$\begin{aligned} P\{X_1 &\geqq c_1; X_2 < c_2\} = [1 - \Phi(c_1)]\,\Phi(c_2) - \Psi(\rho, c_1, c_2) && (1.5)\\ P\{X_1 &< c_1; X_2 \geqq c_2\} = \Phi(c_1)\,[1 - \Phi(c_2)] - \Psi(\rho, c_1, c_2) && (1.6)\\ P\{X_1 &< c_1; X_2 < c_2\} = \Phi(c_1)\Phi(c_2) + \Psi(\rho, c_1, c_2) && (1.7) \end{aligned}$$

Assume now that instead of (1.1) we have

$$\mathcal{D}(X) = \mathcal{N}(\mu, \Sigma) \tag{1.8}$$

where

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

and

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}.$$

Using the fact that

$$\mathcal{D}\left[\begin{pmatrix} (X_1 - \mu_1)/\sigma_1 \\ (X_2 - \mu_2)/\sigma_2 \end{pmatrix}\right] = \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right] \tag{1.9}$$

Date: Modified version (6), May 20, 2013.

where $\rho = \sigma_{12}/\sigma_1\sigma_2$, we obtain

$$\begin{aligned}
P\{X_1 &\geqq c_1; X_2 \geqq c_2\} = \left[1-\Phi\left(\frac{c_1-\mu_1}{\sigma_1}\right)\right]\left[1-\Phi\left(\frac{c_2-\mu_2}{\sigma_2}\right)\right] \\
&\quad +\Psi\left(\rho, \frac{c_1-\mu_1}{\sigma_1}, \frac{c_2-\mu_2}{\sigma_2}\right) &\text{(1.10a)}\\
P\{X_1 &\geqq c_1; X_2 < c_2\} = \left[1-\Phi\left(\frac{c_1-\mu_1}{\sigma_1}\right)\right]\Phi\left(\frac{c_2-\mu_2}{\sigma_2}\right) - \Psi\left(\rho, \frac{c_1-\mu_1}{\sigma_1}, \frac{c_2-\mu_2}{\sigma_2}\right) &\text{(1.10b)}\\
P\{X_1 &< c_1; X_2 \geqq c_2\} = \Phi\left(\frac{c_1-\mu_1}{\sigma_1}\right)\left[1-\Phi\left(\frac{c_2-\mu_2}{\sigma_2}\right)\right] - \Psi\left(\rho, \frac{c_1-\mu_1}{\sigma_1}, \frac{c_2-\mu_2}{\sigma_2}\right) &\text{(1.10c)}\\
P\{X_1 &< c_1; X_2 < c_2\} = \Phi\left(\frac{c_1-\mu_1}{\sigma_1}\right)\Phi\left(\frac{c_2-\mu_2}{\sigma_2}\right) + \Psi\left(\rho, \frac{c_1-\mu_1}{\sigma_1}, \frac{c_2-\mu_2}{\sigma_2}\right) &\text{(1.10d)}
\end{aligned}$$

Consider a three-dimensional random vector $X = (X_1, X_2, X_3)'$ such that

$$\mathcal{D}(X) = \mathcal{N}(0, \Sigma) \tag{1.11}$$

where now

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \sigma_{13} \\ \rho_{12} & 1 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}$$

which may also be written as

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13}\sigma_3 \\ \rho_{12} & 1 & \rho_{23}\sigma_3 \\ \rho_{13}\sigma_3 & \rho_{23}\sigma_3 & \sigma_3^2 \end{pmatrix}.$$

Our purpose is to extend formulae (1.10a) through (1.10d) to conditional probabilities with respect to X_3.

To this end, we observe that

$$\mathcal{D}\left[\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \mid X_3\right] = \mathcal{N}\left[\begin{pmatrix} \frac{\rho_{13}}{\sigma_3}X_3 \\ \frac{\rho_{23}}{\sigma_3}X_3 \end{pmatrix}, \begin{pmatrix} 1-\rho_{13}^2 & \rho_{12}-\rho_{13}\rho_{23} \\ \rho_{12}-\rho_{13}\rho_{23} & 1-\rho_{23}^2 \end{pmatrix}\right]. \tag{1.12}$$

This implies that the conditional probabilities with respect to X_3 that we are seeking to compute are obtained by substituting in (1.10a) through (1.10d) $\mu_1, \mu_2, \sigma_1, \sigma_2$ and ρ for $\frac{\rho_{13}}{\sigma_3}X_3$, $\frac{\rho_{23}}{\sigma_3}X_3$, $\sqrt{1-\rho_{13}^2}$, $\sqrt{1-\rho_{23}^2}$ and $\frac{\rho_{12}-\rho_{13}\rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}}$respectively. We thus obtain,

$$\begin{aligned}
P\{X_1 \geqq c_1; X_2 \geqq c_2 \mid X_3\} &= \left[1-\Phi\left(\frac{c_1-\frac{\rho_{13}}{\sigma_3}X_3}{\sqrt{1-\rho_{13}^2}}\right)\right]\left[1-\Phi\left(\frac{c_2-\frac{\rho_{23}}{\sigma_3}X_3}{\sqrt{1-\rho_{23}^2}}\right)\right] \\
&\quad +\Psi\left(\frac{\rho_{12}-\rho_{13}\rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}}, \frac{c_1-\frac{\rho_{13}}{\sigma_3}X_3}{\sqrt{1-\rho_{13}^2}}, \frac{c_2-\frac{\rho_{23}}{\sigma_3}X_3}{\sqrt{1-\rho_{23}^2}}\right) \qquad (1.13a) \\
P\{X_1 \geqq c_1; X_2 < c_2 \mid X_3\} &= \left[1-\Phi\left(\frac{c_1-\frac{\rho_{13}}{\sigma_3}X_3}{\sqrt{1-\rho_{13}^2}}\right)\right]\Phi\left(\frac{c_2-\frac{\rho_{23}}{\sigma_3}X_3}{\sqrt{1-\rho_{23}^2}}\right) \\
&\quad -\Psi\left(\frac{\rho_{12}-\rho_{13}\rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}}, \frac{c_1-\frac{\rho_{13}}{\sigma_3}X_3}{\sqrt{1-\rho_{13}^2}}, \frac{c_2-\frac{\rho_{23}}{\sigma_3}X_3}{\sqrt{1-\rho_{23}^2}}\right) \qquad (1.13b) \\
P\{X_1 < c_1; X_2 \geqq c_2 \mid X_3\} &= \Phi\left(\frac{c_1-\frac{\rho_{13}}{\sigma_3}X_3}{\sqrt{1-\rho_{13}^2}}\right)\left[1-\Phi\left(\frac{c_2-\frac{\rho_{23}}{\sigma_3}X_3}{\sqrt{1-\rho_{23}^2}}\right)\right] \\
&\quad -\Psi\left(\frac{\rho_{12}-\rho_{13}\rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}}, \frac{c_1-\frac{\rho_{13}}{\sigma_3}X_3}{\sqrt{1-\rho_{13}^2}}, \frac{c_2-\frac{\rho_{23}}{\sigma_3}X_3}{\sqrt{1-\rho_{23}^2}}\right) \qquad (1.13c) \\
P\{X_1 < c_1; X_2 < c_2 \mid X_3\} &= \Phi\left(\frac{c_1-\frac{\rho_{13}}{\sigma_3}X_3}{\sqrt{1-\rho_{13}^2}}\right)\Phi\left(\frac{c_2-\frac{\rho_{23}}{\sigma_3}X_3}{\sqrt{1-\rho_{23}^2}}\right) \\
&\quad +\Psi\left(\frac{\rho_{12}-\rho_{13}\rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}}, \frac{c_1-\frac{\rho_{13}}{\sigma_3}X_3}{\sqrt{1-\rho_{13}^2}}, \frac{c_2-\frac{\rho_{23}}{\sigma_3}X_3}{\sqrt{1-\rho_{23}^2}}\right) \qquad (1.13d)
\end{aligned}$$

Of course, the unconditional probabilities are obtained by taking the expectation with respect to X_3 of the corresponding conditional probabilities.

2. Estimation of the model

Our aim is to estimate the simultaneous three-equation model with mixed latent and observed variables.

Define, omitting the index n, $n = 1, \dots, N$, for the individual observations.

$$\left\{\begin{array}{l} y_1^* : \text{ propensity to consult} \\ y_2^* : \text{ propensity to be prescribed a drug} \\ y_3^* : \text{ health status (latent, unobserved)} \end{array}\right. \qquad (2.14)$$

The observed variable $y_1$is a count variable for the number of visits:

The distribution of y_1 conditional on a vector of exogenous variable x_1 and an unobserved disturbance term u_1^0 is a Poisson distribution such that

$$\begin{aligned} \mathcal{D}(y_1 \quad | \quad & x_1, u_1^0) = \text{Poisson } (\lambda^0), \qquad (2.15) \\ \text{where } \lambda^0 \quad = \quad & \exp(x_1'\beta_1^0 + u_1^0). \end{aligned}$$

Consider the simultaneous-equations model with mixed latent, dichotomous and count variables

$$\left\{\begin{array}{l} y_1 \text{ such that } \mathcal{D}(y_1 \mid x_1, u_1^0) \text{ is the above mentioned Poisson distribution} \\ y_2^* = x_2'\beta_2^0 + \alpha_{21}^0 y_1 + u_2^0 \\ y_3^* = x_3'\beta_3^0 + \alpha_{31}^0 y_1 + \alpha_{32}^0 y_2 + u_3^0 \end{array}\right. \qquad (2.16)$$

We assume that

$$\mathcal{D}\left(u^0\right) = \mathcal{N}(0, \Sigma^0), \qquad (2.17)$$

where

$$u^0 = \begin{pmatrix} u_1^0 \\ u_2^0 \\ u_3^0 \end{pmatrix} \tag{2.18}$$

and

$$\Sigma^0 = \begin{pmatrix} \sigma_1^{02} & \rho_{12}^0\sigma_1^0 & \rho_{13}^0\sigma_1^0 \\ \rho_{12}^0\sigma_1^0 & 1 & \rho_{23}^0 \\ \rho_{13}^0\sigma_1^0 & \rho_{23}^0 & 1 \end{pmatrix} \tag{2.19}$$

Under this assumption, the unconditional distribution of $y_1$is the so-called *Poisson log-normal distribution* (hereafter PLN) and not the more familiar negative-binomial distribution. The PLN distribution is obtained as a mixed Poisson distribution if a log-normal distribution, $\mathcal{LN}(\mu^0, \sigma^{02})$ is used as a mixing density.

We observe

$$y_2 = \begin{cases} 1 \text{ if } y_2^* \geq 0 \\ 0 \text{ otherwise} \end{cases}, \tag{2.20}$$

and the polychotomous ordered-response variable y_3 such that

$$y_3 = \begin{cases} 1 & \text{if } c_0 \leq y_3^* < c_1 \\ 2 & \text{if } c_1 \leq y_3^* < c_2 \\ 3 & \text{if } c_2 \leq y_3^* < c_3 \\ 4 & \text{if } c_3 \leq y_3^* < c_4 \\ 5 & \text{if } c_4 \leq y_3^* < c_5 \end{cases} \tag{2.21}$$

where the set of constants c's are such that $c_0 = -\infty$, $c_5 = +\infty$ and $c_0 < c_1 < c_2 < c_3 < c_4 < c_5$. We say that y_3^* belongs to the jth category if

$$c_{j-1} \leq y_3^* < c_j \quad (j = 1, 2, \ldots, J) \text{ with } J = 5.$$

2.1. **Intermediary step 1: Estimation of the Poisson log-normal distribution.** Define $d\varphi_{\sigma_1^0}(u_1^0)$ as

$$d\varphi_{\sigma_1^0}(u_1^0) = \frac{1}{\sigma_1^0}\varphi\left(\frac{u_1^0}{\sigma_1^0}\right) du_1^0.$$

In the case where we only seek the marginal distribution of the Poisson log-normal, we have to compute the probabilities:

$$\begin{aligned} P(y_1 &= k) = \int_{-\infty}^{+\infty} P(y_1 = k \mid u_1^0) d\varphi_{\sigma_1^0}(u_1^0) \\ &= \frac{1}{\sqrt{2\pi}\sigma_1^0}\frac{1}{k!}\int_{-\infty}^{+\infty} \exp\left[-\left(\mu^0 \exp u_1^0\right)\right]\left(\mu^0 \exp u_1^0\right)^k \exp\left(-\frac{u_1^{02}}{2\sigma_1^{02}}\right) du_1^0, \end{aligned}$$

where

$$\mu^0 = \exp(x_1'\beta_1^0)$$

We may thus write

$$\begin{aligned} P(y_1 &= k) = \frac{1}{\sqrt{2\pi}\sigma_1^0}\frac{1}{k!}\mu^{0k}\int_{-\infty}^{+\infty} \exp\left[-\left(\mu^0 \exp u_1^0\right)\right] \exp k u_1^0 \exp\left(-\frac{u_1^{02}}{2\sigma_1^{02}}\right) du_1^0 \\ &= \frac{1}{\sqrt{2\pi}\sigma_1^0}\frac{1}{k!}\mu^{0k}\int_{-\infty}^{+\infty} \exp\left[-\left(\mu^0 \exp u_1^0\right) + k u_1^0\right] \exp\left(-\frac{u_1^{02}}{2\sigma_1^{02}}\right) du_1^0 \end{aligned} \tag{2.22}$$

The estimation of this model has been considered among others by Greene (1977). He suggested to carry-out the computation of these probabilities using the Gauss-Hermite quadrature. Miranda

(2004) suggested a STATA code for this Gauss-Hermite numerical computation. Specifically, let in (2.22) $v^0 = u_1^0/\sqrt{2}\sigma_1^0$. With this change of variables (2.22) may be re-expressed as

$$P(y_1 = k) = \frac{1}{\sqrt{\pi}}\frac{1}{k!}\mu^{0k}\int_{-\infty}^{+\infty}\exp\left[-\left(\mu^0\exp\sqrt{2}\sigma_1^0 v^0\right) + k\sqrt{2}\sigma_1^0 v^0\right]\exp\left(-v^{02}\right)dv^0 \tag{2.23}$$

As suggested by Greene (1997), the value of the probalilities $P(y_1 = k)$ can be approximated by using Gauss–Hermite quadrature for the integration and compute $P^*(y_1 = k)$ where

$$P^*(y_1 = k) = \frac{1}{\sqrt{\pi}}\frac{1}{k!}\mu^{0k}\sum_{l=1}^{L} w_l \exp\left[-\left(\mu^0\exp\sqrt{2}\sigma_1^0 v_l^0\right) + k\sqrt{2}\sigma_1^0 v_l^0\right] \tag{2.24}$$

where L is the number of sample points to use for the approximation. The v_l^0 are the roots of the Hermite polynomial $H_L(x)$ $(l = 1, 2, \ldots, L)$ and the associated weights w_l are given by

$$w_l = \frac{2^{L-1}L!\sqrt{\pi}}{L^2\left[H_{L-1}(v_l^0)\right]^2}$$

2.2. **Intermediary step 2: Estimation for first two two-equation model.** We now consider as an intermediary step the estimation of the simultaneous equation model made of the first two equation considered above. Explicitly, we consider the following two-equation model

$$\left\{\begin{array}{l} y_1 \text{ such that } \mathcal{D}(y_1 \mid x_1, u_1^0) \text{ is the above mentioned Poisson distribution} \\ y_2^* = x_2'\beta_2^0 + \alpha_{21}^0 y_1 + u_2^0 \end{array}\right. \tag{2.25}$$

where we observe

$$y_2 = \left\{\begin{array}{l} 1 \text{ if } y_2^* \geq 0 \\ 0 \text{ otherwise} \end{array}\right. . \tag{2.26}$$

We seek to find the expression of $P(y_1 = k, y_2 = i)$ for $k = 0, 1, \ldots$ and $i = 0, 1$.

We have

$$P(y_1 = k, y_2 = i) = \int_{-\infty}^{+\infty} P(y_2 = i \mid y_1 = k, u_1^0)P(y_1 = k \mid u_1^0)d\varphi_{\sigma_1^0}(u_1^0) \tag{2.27}$$

We need to derive the expression for the conditional probabilities $P(y_2 = i \mid y_1 = k, u_1^0)$.

To this end, we use the fact that

$$\mathcal{D}\left(u_2^0 \mid u_1^0\right) = \mathcal{N}\left(\frac{\rho_{12}^0}{\sigma_1^0}u_1^0, 1 - \rho_{12}^{02}\right). \tag{2.28}$$

We have, for $i = 1$,

$$\begin{aligned} P(y_2 &= 1 \mid y_1 = k, u_1^0) = P(y_2^* \geq 0 \mid y_1 = k, u_1^0) \\ &= P(x_2'\beta_2^0 + \alpha_{21}^0 y_1 + u_2^0 \geq 0 \mid y_1 = k, u_1^0) \\ &= P(u_2^0 \geq -x_2'\beta_2^0 - \alpha_{21}^0 y_1 \mid y_1 = k, u_1^0) \\ &= P(\frac{u_2^0 - \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}} \geq -\frac{x_2'\beta_2^0 + \alpha_{21}^0 y_1 + \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}} \mid y_1 = k, u_1^0) \end{aligned}$$

so that, using (2.28), we obtain

$$\begin{aligned} P(y_2 &= 1 \mid y_1 = k, u_1^0) = 1 - \Phi\left(-\frac{x_2'\beta_2^0 + \alpha_{21}^0 k + \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right) \\ &= \Phi\left(\frac{x_2'\beta_2^0 + \alpha_{21}^0 k + \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right) \end{aligned} \tag{2.29}$$

We also have, for $i = 0$,

$$\begin{aligned} P(y_2 &= 0 \mid y_1 = k, u_1^0) = 1 - P(y_2 = 1 \mid y_1 = k, u_1^0) \\ &= 1 - \Phi\left(\frac{x_2'\beta_2^0 + \alpha_{21}^0 k + \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right) \end{aligned} \tag{2.30}$$

It follows that

$$\begin{aligned} P(y_1 &= k, y_2 = 1) = \int_{-\infty}^{+\infty} P(y_2 = 1 \mid y_1 = k, u_1^0) P(y_1 = k \mid u_1^0) d\varphi_{\sigma_1^0}(u_1^0) \\ &= \frac{1}{\sqrt{2\pi}\sigma_1^0}\frac{1}{k!}\mu^{0k}\int_{-\infty}^{+\infty}\Phi\left(\frac{x_2'\beta_2^0 + \alpha_{21}^0 k + \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right)\exp\left[-\left(\mu^0\exp u_1^0\right) + k u_1^0\right]\exp\left(-\frac{u_1^{02}}{2\sigma_1^{02}}\right)du_1^0 \end{aligned} \tag{2.31}$$

where, as before,

$$\mu^0 = \exp(x_1'\beta_1^0)$$

and

$$\begin{aligned} P(y_1 &= k, y_2 = 0) = \int_{-\infty}^{+\infty} P(y_2 = 0 \mid y_1 = k, u_1^0) P(y_1 = k \mid u_1^0) d\varphi_{\sigma_1^0}(u_1^0) \\ &= \frac{1}{\sqrt{2\pi}\sigma_1^0}\frac{1}{k!}\mu^{0k}\int_{-\infty}^{+\infty}\left[1-\Phi\left(\frac{x_2'\beta_2^0 + \alpha_{21}^0 k + \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right)\right]\exp\left[-\left(\mu^0\exp u_1^0\right) + k u_1^0\right]\exp\left(-\frac{u_1^{02}}{2\sigma_1^{02}}\right)du_1^0 \end{aligned} \tag{2.32}$$

Note that, from the computational point of view, that

$$P(y_1 = k, y_2 = 0) = P(y_1 = k) - P(y_1 = k, y_2 = 1), \tag{2.33}$$

which possibly provides a simpler alternative way to compute $P(y_1 = k, y_2 = 0)$ once $P(y_1 = k)$ and $P(y_1 = k, y_2 = 1)$ have been computed.

The log-likelihood function for a sample with N observations may be written as

$$\mathcal{L}_N = \sum_{n=1}^{N}\sum_{i=0}^{1}\sum_{k} z_{nki}\log P_{nki} \tag{2.34}$$

where

$$z_{nki} = \begin{cases} 1 & \text{if } y_{n1} = k \text{ and } y_{n2} = i \\ 0 & \text{otherwise} \end{cases} \tag{2.35}$$

and P_{nki} is defined as

$$P_{nki} = P\{z_{nki} = 1\} \tag{2.36}$$

that is to say

$$\begin{aligned} P_{nki} &= P(y_{n1}=k, y_{n2}=i) \\ &= \int_{-\infty}^{+\infty} P(y_{n2}=i \mid y_{n1}=k, u_{n1}) P(y_{n1}=k \mid u_{n1}) d\varphi_{\sigma_1}(u_{n1}) \end{aligned} \tag{2.37}$$

Explicitly, we have

$$P_{nk1} = \frac{1}{\sqrt{2\pi}\sigma_1} \frac{1}{k!} \mu_n^k \int_{-\infty}^{+\infty} \Phi\left(\frac{x'_{n2}\beta_2 + \alpha_{21}k + \frac{\rho_{12}}{\sigma_1}u_{n1}}{\sqrt{1-\rho_{12}^2}} \right) \exp\left[-\left(\mu_n \exp u_{n1}\right) + k u_{n1} \right] \exp\left(-\frac{u_{n1}^2}{2\sigma_1^2} \right) du_{n1} \tag{2.38}$$

where

$$\mu_n = \exp(x'_{n1}\beta_1) \tag{2.39}$$

and

$$P_{nk0} = \frac{1}{\sqrt{2\pi}\sigma_1} \frac{1}{k!} \mu_n^k \int_{-\infty}^{+\infty} \left[1 - \Phi\left(\frac{x'_{n2}\beta_2 + \alpha_{21}k + \frac{\rho_{12}}{\sigma_1}u_{n1}}{\sqrt{1-\rho_{12}^2}} \right) \right] \exp\left[-\left(\mu_n \exp u_{n1}\right) + k u_{n1} \right] \exp\left(-\frac{u_{n1}^2}{2\sigma_1^2} \right) du_{n1} \tag{2.40}$$

In view of possible computation of the probabilities P_{nk1} and P_{nk0} using the Gauss-Hermite quadrature, we may make the change of variable $v_n = u_{n1}/\sqrt{2}\sigma_1$ in the expressions of these probabilities and obtain

$$P_{nk1} = \frac{1}{\sqrt{\pi}} \frac{1}{k!} \mu_n^k \int_{-\infty}^{+\infty} \Phi\left(\frac{x'_{n2}\beta_2 + \alpha_{21}k + \rho_{12}\sqrt{2}v_n}{\sqrt{1-\rho_{12}^2}} \right) \exp\left[-\left(\mu_n \exp \sqrt{2}\sigma_1 v_n\right) + k\sqrt{2}\sigma_1 v_n \right] \exp\left(-v_n^2\right) dv_n \tag{2.41}$$

and

$$P_{nk0} = \frac{1}{\sqrt{\pi}} \frac{1}{k!} \mu_n^k \int_{-\infty}^{+\infty} \left[1 - \Phi\left(\frac{x'_{n2}\beta_2 + \alpha_{21}k + \rho_{12}\sqrt{2}v_n}{\sqrt{1-\rho_{12}^2}} \right) \right] \exp\left[-\left(\mu_n \exp \sqrt{2}\sigma_1 v_n\right) + k\sqrt{2}\sigma_1 v_n \right] \exp\left(-v_n^2\right) d \tag{2.42}$$

2.3. **The complete three-equations model.** Consider now the complete three-equations model described above. We seek to find the expression of $P(y_1=k, y_2=i, y_3=j)$ for $k=0,1,\ldots$, $i=0,1$ and $j=1,\ldots,5$.

We have

$$P(y_1=k, y_2=i, y_3=j) = \int_{-\infty}^{+\infty} P(y_2=i, y_3=j \mid y_1=k, u_1^0) P(y_1=k \mid u_1^0) d\varphi_{\sigma_1^0}(u_1^0) \tag{2.43}$$

where, as before, $d\varphi_{\sigma_1^0}(u_1^0)$ is defined as

$$d\varphi_{\sigma_1^0}(u_1^0) = \frac{1}{\sigma_1^0}\varphi\left(\frac{u_1^0}{\sigma_1^0}\right) du_1^0.$$

2.3.1. *Derivation of the probabilities.* We need to derive the expression for the conditional probabilities $P(y_2=i, y_3=j \mid y_1=k, u_1^0)$.

We use the fact that

$$\mathcal{D}\left[\begin{pmatrix} u_2^0 \\ u_3^0 \end{pmatrix} \mid u_1^0 \right] = \mathcal{N}\left[\begin{pmatrix} \frac{\rho_{12}^0}{\sigma_1^0}u_1^0 \\ \frac{\rho_{13}^0}{\sigma_1^0}u_1^0 \end{pmatrix}, \begin{pmatrix} 1-\rho_{12}^{02} & \rho_{23}^0 - \rho_{12}^0\rho_{13}^0 \\ \rho_{23}^0 - \rho_{12}^0\rho_{13}^0 & 1-\rho_{13}^{02} \end{pmatrix} \right]. \tag{2.44}$$

For $i=0$ we have,

$$
\begin{aligned}
P\{y_2 &= 0, y_3 = j \mid y_1 = k, u_1^0\} = P\{y_2^* < 0, c_{j-1} \leq y_3^* < c_j \mid y_1 = k, u_1^0\} \\
&= P\{y_2^* < 0, y_3^* < c_j \mid y_1 = k, u_1^0\} - P\{y_2^* < 0, y_3^* < c_{j-1} \mid y_1 = k, u_1^0\} \\
&= P\{x_2'\beta_2^0 + \alpha_{21}^0 k + u_2^0 < 0, x_3'\beta_3^0 + \alpha_{31}^0 k + u_3^0 < c_j \mid u_1^0\} \\
-P\{x_2'\beta_2^0 + \alpha_{21}^0 k + u_2^0 &< 0, x_3'\beta_3^0 + \alpha_{31}^0 k + u_3^0 < c_{j-1} \mid u_1^0\} \\
&= P\{u_2^0 < -x_2'\beta_2^0 - \alpha_{21}^0 k, u_3^0 < c_j - x_3'\beta_3^0 - \alpha_{31}^0 k \mid u_1^0\} \\
-P\{u_2^0 &< -x_2'\beta_2^0 - \alpha_{21}^0 k, u_3^0 < c_{j-1} - x_3'\beta_3^0 - \alpha_{31}^0 k \mid u_1^0\},
\end{aligned}
$$

which, using (1.13d), may be written as

$$
\begin{aligned}
P\{y_2 = 0, y_3 = j \mid y_1 = k, u_1^0\} = \Phi\left(\frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right) \Bigg[\Phi\left(\frac{c_j - x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right) \\
-\Phi\left(\frac{c_{j-1} - x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\Bigg] \\
+\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{c_j - x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right) \\
-\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{c_{j-1} - x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)
\end{aligned}
\tag{2.45}
$$

Similarly, for $i = 1$ we have,

$$
\begin{aligned}
P\{y_2 &= 1, y_3 = j \mid y_1 = k, u_1^0\} = P\{y_2^* \geq 0, c_{j-1} \leq y_3^* < c_j \mid y_1 = k, u_1^0\} \\
&= P\{y_2^* \geq 0, y_3^* < c_j \mid y_1 = k, u_1^0\} - P\{y_2^* \geq 0, y_3^* < c_{j-1} \mid y_1 = k, u_1^0\} \\
&= P\{x_2'\beta_2^0 + \alpha_{21}^0 k + u_2^0 \geq 0, x_3'\beta_3^0 + \alpha_{31}^0 y_1 + \alpha_{32}^0 + u_3^0 < c_j \mid u_1^0\} \\
-P\{x_2'\beta_2^0 + \alpha_{21}^0 k + u_2^0 &\geq 0, x_3'\beta_3^0 + \alpha_{31}^0 y_1 + \alpha_{32}^0 + u_3^0 < c_{j-1} \mid u_1^0\} \\
&= P\{u_2^0 \geq -x_2'\beta_2^0 - \alpha_{21}^0 k, u_3^0 < c_j - x_3'\beta_3^0 - \alpha_{31}^0 k - \alpha_{32}^0 \mid u_1^0\} \\
-P\{u_2^0 &\geq -x_2'\beta_2^0 - \alpha_{21}^0 k, u_2^0 < c_{j-1} - x_3'\beta_3^0 - \alpha_{31}^0 k - \alpha_{32}^0 \mid u_1^0\},
\end{aligned}
$$

which, using (1.13b), may be written as

$$
\begin{aligned}
P\{y_2 = 1, y_3 = j \mid y_1 = k, u_1^0\} = \left[1 - \Phi\left(\frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right)\right] \times \\
\times \left[\Phi\left(\frac{c_j - x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right) - \Phi\left(\frac{c_{j-1} - x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right] \\
-\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{c_j - x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right) \\
+\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{c_{j-1} - x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)
\end{aligned}
\tag{2.46}
$$

Note that

$$P\{y_2=1,y_3=j\mid y_1=k,u_1^0\}=\left[\Phi\left(\frac{c_j-x_3'\beta_3^0-\alpha_{31}^0k-\frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right.$$
$$\left.-\Phi\left(\frac{c_{j-1}-x_3'\beta_3^0-\alpha_{31}^0k-\frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right]-P\{y_2=0,y_3=j\mid y_1=k,u_1^0\},$$

which allows to verify that, as one should expect,

$$\begin{aligned}P\{y_3 &= j\mid y_1=k,u_1^0\}=P\{y_2=0,y_3=j\mid y_1=k,u_1^0\}+P\{y_2=1,y_3=j\mid y_1=k,u_1^0\}\\ &= \left[\Phi\left(\frac{c_j-x_3'\beta_3^0-\alpha_{31}^0k-\frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)-\Phi\left(\frac{c_{j-1}-x_3'\beta_3^0-0688_{31}^0k-\frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right]\end{aligned}$$

Equations (2.45) and (2.46) allow us to compute all the probabilities $P(y_1=k,y_2=i,y_3=j)$ for $k=0,1,\ldots,$ $i=0,1$ and $j=1,\ldots,5$ using (2.43).

2.3.2. *The log-likelihood function.* The log-likelihood function for a sample with N observation may be written as

$$\mathcal{L}_N=\sum_{n=1}^{N}\sum_{k=0}^{\infty}\sum_{i=0}^{1}\sum_{j=1}^{5}z_{nkij}\log P_{nkij} \tag{2.47}$$

where

$$z_{nkij}=\begin{cases}1 & \text{if } y_{n1}=k,\ y_{n2}=i \text{ and } y_{n3}=j \quad (k=0,1,\ldots,i=0,1 \text{ and } j=1,\ldots,5)\\ 0 \text{ otherwise}\end{cases} \tag{2.48}$$

and P_{nij} is defined as

$$P_{nkij}=P\{z_{nkij}=1\} \tag{2.49}$$

that is to say

$$\begin{aligned}P_{nkij} &= P(y_{n1}=k,y_{n2}=i,y_{n3}=j)\\ &= \int_{-\infty}^{+\infty}P(y_{n2}=i,y_{n3}=j\mid y_{n1}=k,u_{n1})P(y_{n1}=k\mid u_{n1})d\varphi_{\sigma_1}(u_{n1})\end{aligned} \tag{2.50}$$

where $d\varphi_{\sigma_1^0}(u_1^0)$ is defined as

$$d\varphi_{\sigma_1}(u_{n1})=\frac{1}{\sigma_1}\varphi\left(\frac{u_{n1}}{\sigma_1}\right)du_{n1}.$$

Explicitly,

$$P_{nkij}=\frac{1}{\sqrt{2\pi}\sigma_1}\int_{-\infty}^{+\infty}P(y_{n2}=i,y_{n3}=j\mid y_{n1}=k,u_{n1})P(y_{n1}=k\mid u_{n1})\exp\left(-\frac{u_{n1}^2}{2\sigma_1^2}\right)du_{n1}. \tag{2.51}$$

From a numerical point of view, it is convenient to make the change of variable $v_n=u_{n1}/\sqrt{2}\sigma_1$ so that (2.51) may be written as

$$P_{nkij}=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}P(y_{n2}=i,y_{n3}=j\mid y_{n1}=k,\sqrt{2}\sigma_1v_n)P(y_{n1}=k\mid\sqrt{2}\sigma_1v_n)\exp\left(-v_n^2\right)dv_n. \tag{2.52}$$

The probabilities $P(y_{n2}=i,y_{n3}=j\mid y_{n1}=k,\sqrt{2}\sigma_1v_n)$ can be obtained from (2.45) and (2.46) by making the change in variable $u_1^0=\sqrt{2}\sigma_1^0v^0$.

In order to derive the expression for the probability $P(y_{n1} = k \mid \sqrt{2}\sigma_1 v_n)$, it is convenient to define μ_n as

$$\mu_n = \exp(x'_{n1}\beta_1)$$

so that, similarly as in (**??**) we may write $P(y_{n1} = k \mid \sqrt{2}\sigma_1 v_n)$ as

$$P(y_{n1} = k \mid \sqrt{2}\sigma_1 v_n) = \frac{1}{k!}\mu_n^k \exp\left[-\left(\mu_n \exp\sqrt{2}\sigma_1 v_n\right) + k\sqrt{2}\sigma_1 v_n\right] \tag{2.53}$$

The value of the probalilities P_{nkij} in (2.76) can be approximated by using Gauss–Hermite quadrature for the integration and compute P^*_{nkij} where

$$P^*_{nkij} = \frac{1}{\sqrt{\pi}}\sum_{l=1}^{L} w_{nl} P(y_{n2} = i, y_{n3} = j \mid y_{n1} = k, \sqrt{2}\sigma_1 v_{nl}) P(y_{n1} = k \mid \sqrt{2}\sigma_1 v_{nl}) \tag{2.54}$$

where, as before, L is the number of sample points to use for the approximation. The v_{nl} are the roots of the Hermite polynomial $H_L(x)$ ($l = 1, 2, \ldots, L$) and the associated weights w_l are given by

$$w_{nl} = \frac{2^{L-1}L!\sqrt{\pi}}{L^2\left[H_{L-1}(v_{nl})\right]^2}$$

2.3.3. *The case of a dichotomous y_3.* Consider the simplest case where y_3 is dichomotous,

$$y_3 = \begin{cases} 1 \text{ if } y_3^* \geq 0 \\ 0 \text{ otherwise} \end{cases}, \tag{2.55}$$

For $i = 0$ and $y_3 = 1$ we have,

$$\begin{aligned} P\{y_2 &= 0, y_3 = 1 \mid y_1 = k, u_1^0\} = P\{y_2^* < 0, y_3^* \geq 0 \mid y_1 = k, u_1^0\} \\ &= P\{x'_2\beta_2^0 + \alpha_{21}^0 y_1 + u_2^0 < 0, x'_3\beta_3^0 + \alpha_{31}^0 y_1 + \alpha_{32}^0 y_2 + u_3^0 \geq 0 \mid y_1 = k, u_1^0\} \\ &= P\{x'_2\beta_2^0 + \alpha_{21}^0 k + u_2^0 < 0, x'_3\beta_3^0 + \alpha_{31}^0 k + u_3^0 \geq 0 \mid u_1^0\} \\ &= P\{u_2^0 < -x'_2\beta_2^0 - \alpha_{21}^0 k, u_3^0 \geq -x'_3\beta_3^0 - \alpha_{31}^0 k \mid u_1^0\} \end{aligned} \tag{2.56}$$

which, using (1.13c), may be written as

$$\begin{aligned} P\{y_2 &= 0, y_3 = 1 \mid y_1 = k, u_1^0\} = \Phi\left(\frac{-x'_2\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right)\left[1 - \Phi\left(\frac{-x'_3\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right] \\ &-\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x'_2\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{-x'_3\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right) \end{aligned} \tag{2.57}$$

Therefore, the unconditional probability $P\{y_1 = k, y_2 = 0, y_3 = 1\}$ is equal to

$$P\{y_1 = k, y_2 = 0, y_3 = 1\} = \int_{-\infty}^{+\infty} P\{y_2 = 0, y_3 = 1 \mid y_1 = k, u_1^0\} P\left\{y_1 = k \mid u_1^0\right\} d\varphi_{\sigma_1^0}(u_1^0),$$

and thus,

$$\begin{aligned}
P\{y_1 \;&=\; k, y_2 = 0, y_3 = 1\} = \frac{1}{\sqrt{2\pi}\sigma_1^0}\frac{1}{k!}\mu^{0k}\int_{-\infty}^{+\infty}\left\{\Phi\left(\frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right)\right. \\
&\times\left[1 - \Phi\left(\frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right] \qquad (2.58)\\
&\left. -\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right\} \\
&\times \exp\left[-\left(\mu^0 \exp u_1^0\right) + k u_1^0\right]\exp\left(-\frac{u_1^{02}}{2\sigma_1^{02}}\right)du_1^0 \qquad (2.59)
\end{aligned}$$

Similarly, for $i = 0$ and $y_3 = 0$ we have,

$$\begin{aligned}
P\{y_2 \;&=\; 0, y_3 = 0 \mid y_1 = k, u_1^0\} = P\{y_2^* < 0, y_3^* < 0 \mid y_1 = k, u_1^0\} \\
&= P\{x_2'\beta_2^0 + \alpha_{21}^0 y_1 + u_2^0 < 0, x_3'\beta_3^0 + \alpha_{31}^0 y_1 + \alpha_{32}^0 y_2 + u_3^0 < 0 \mid y_1 = k, u_1^0\} \\
&= P\{x_2'\beta_2^0 + \alpha_{21}^0 k + u_2^0 < 0, x_3'\beta_3^0 + \alpha_{31}^0 k + u_3^0 \geq 0 \mid u_1^0\} \\
&= P\{u_2^0 < -x_2'\beta_2^0 - \alpha_{21}^0 k, u_3^0 < -x_3'\beta_3^0 - \alpha_{31}^0 k \mid u_1^0\} \qquad (2.60)
\end{aligned}$$

which, using (1.13d), may be written as

$$\begin{aligned}
P\{y_2 \;&=\; 0, y_3 = 0 \mid y_1 = k, u_1^0\} = \Phi\left(\frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right)\left[\Phi\left(\frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right] \\
&+\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right) \qquad (2.61)
\end{aligned}$$

Therefore, the unconditional probability $P\{y_1 = k, y_2 = 0, y_3 = 0\}$ is equal to

$$P\{y_1 = k, y_2 = 0, y_3 = 0\} = \int_{-\infty}^{+\infty} P\{y_2 = 0, y_3 = 0 \mid y_1 = k, u_1^0\} P\left\{y_1 = k \mid u_1^0\right\} d\varphi_{\sigma_1^0}(u_1^0),$$

and thus,

$$\begin{aligned}
P\{y_1 \;&=\; k, y_2 = 0, y_3 = 0\} = \frac{1}{\sqrt{2\pi}\sigma_1^0}\frac{1}{k!}\mu^{0k}\int_{-\infty}^{+\infty}\left\{\Phi\left(\frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right)\left[\Phi\left(\frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right.\right.\right. \\
&\left.+\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right\} \\
&\times \exp\left[-\left(\mu^0 \exp u_1^0\right) + k u_1^0\right]\exp\left(-\frac{u_1^{02}}{2\sigma_1^{02}}\right)du_1^0 \qquad (2
\end{aligned}$$

For $i = 1$ and $y_3 = 1$ we have,

$$
\begin{aligned}
P\{y_2 &= 1, y_3 = 1 \mid y_1 = k, u_1^0\} = P\{y_2^* \geq 0, y_3^* \geq 0 \mid y_1 = k, u_1^0\} \\
&= P\{x_2'\beta_2^0 + \alpha_{21}^0 y_1 + u_2^0 \geq 0, x_3'\beta_3^0 + \alpha_{31}^0 y_1 + \alpha_{32}^0 + u_3^0 \geq 0 \mid y_1 = k, u_1^0\} \\
&= P\{x_2'\beta_2^0 + \alpha_{21}^0 k + u_2^0 \geq 0, x_3'\beta_3^0 + \alpha_{31}^0 k + \alpha_{32}^0 + u_3^0 \geq 0 \mid u_1^0\} \\
&= P\{u_2^0 \geq -x_2'\beta_2^0 - \alpha_{21}^0 k, u_3^0 \geq -x_3'\beta_3^0 - \alpha_{31}^0 k - \alpha_{32}^0 \mid u_1^0\} \qquad (2.63)
\end{aligned}
$$

which, using (1.13a), may be written as

$$
\begin{aligned}
P\{y_2 &= 1, y_3 = 1 \mid y_1 = k, u_1^0\} = \left[1 - \Phi\left(\frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1 - \rho_{12}^{02}}}\right)\right]\left[1 - \Phi\left(\frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \alpha_{32}^0 - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1 - \rho_{13}^{02}}}\right)\right] \\
&+\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1 - \rho_{12}^{02})(1 - \rho_{13}^{02})}}, \frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1 - \rho_{12}^{02}}}, \frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \alpha_{32}^0 - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1 - \rho_{13}^{02}}}\right) \qquad (2
\end{aligned}
$$

Therefore, the unconditional probability $P\{y_1 = k, y_2 = 1, y_3 = 1\}$ is equal to

$$
P\{y_1 = k, y_2 = 1, y_3 = 1\} = \int_{-\infty}^{+\infty} P\{y_2 = 1, y_3 = 1 \mid y_1 = k, u_1^0\} P\left\{y_1 = k \mid u_1^0\right\} d\varphi_{\sigma_1^0}(u_1^0),
$$

and thus,

$$
\begin{aligned}
P\{y_1 &= k, y_2 = 1, y_3 = 1\} = \frac{1}{\sqrt{2\pi}\sigma_1^0}\frac{1}{k!}\mu^{0k}\int_{-\infty}^{+\infty}\left\{\left[1 - \Phi\left(\frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1 - \rho_{12}^{02}}}\right)\right]\right. \\
&\times\left[1 - \Phi\left(\frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \alpha_{32}^0 - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1 - \rho_{13}^{02}}}\right)\right] \\
&\left.+\Psi\left(\frac{\rho_{23}^0 - \rho_{12}^0\rho_{13}^0}{\sqrt{(1 - \rho_{12}^{02})(1 - \rho_{13}^{02})}}, \frac{-x_2'\beta_2^0 - \alpha_{21}^0 k - \frac{\rho_{12}^0}{\sigma_1^0} u_1^0}{\sqrt{1 - \rho_{12}^{02}}}, \frac{-x_3'\beta_3^0 - \alpha_{31}^0 k - \alpha_{32}^0 - \frac{\rho_{13}^0}{\sigma_1^0} u_1^0}{\sqrt{1 - \rho_{13}^{02}}}\right)\right\} \\
&\times \exp\left[-\left(\mu^0 \exp u_1^0\right) + k u_1^0\right] \exp\left(-\frac{u_1^{02}}{2\sigma_1^{02}}\right) du_1^0 \qquad (2.65)
\end{aligned}
$$

Finally, for $i = 1$ and $y_3 = 0$ we have,

$$
\begin{aligned}
P\{y_2 &= 1, y_3 = 0 \mid y_1 = k, u_1^0\} = P\{y_2^* \geq 0, y_3^* < 0 \mid y_1 = k, u_1^0\} \\
&= P\{x_2'\beta_2^0 + \alpha_{21}^0 y_1 + u_2^0 \geq 0, x_3'\beta_3^0 + \alpha_{31}^0 y_1 + \alpha_{32}^0 + u_3^0 < 0 \mid y_1 = k, u_1^0\} \\
&= P\{x_2'\beta_2^0 + \alpha_{21}^0 k + u_2^0 \geq 0, x_3'\beta_3^0 + \alpha_{31}^0 k + +\alpha_{32}^0 + u_3^0 \geq 0 \mid u_1^0\} \\
&= P\{u_2^0 \geq -x_2'\beta_2^0 - \alpha_{21}^0 k, u_3^0 < -x_3'\beta_3^0 - \alpha_{31}^0 k - \alpha_{32}^0 \mid u_1^0\} \qquad (2.66)
\end{aligned}
$$

which, using (1.13b), may be written as

$$P\{y_2 = 1, y_3 = 0 \mid y_1 = k, u_1^0\} = \left[1-\Phi\left(\frac{-x_2'\beta_2^0-\alpha_{21}^0 k-\frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right)\right]\left[\Phi\left(\frac{-x_3'\beta_3^0-\alpha_{31}^0 k-\alpha_{32}^0-\frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right.$$
$$-\Psi\left(\frac{\rho_{23}^0-\rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x_2'\beta_2^0-\alpha_{21}^0 k-\frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{-x_3'\beta_3^0-\alpha_{31}^0 k-\alpha_{32}^0-\frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right) \tag{2.67}$$

Therefore, the unconditional probability $P\{y_1 = k, y_2 = 1, y_3 = 0\}$ is equal to

$$P\{y_1 = k, y_2 = 1, y_3 = 0\} = \int_{-\infty}^{+\infty} P\{y_2 = 1, y_3 = 0 \mid y_1 = k, u_1^0\} P\left\{y_1 = k \mid u_1^0\right\} d\varphi_{\sigma_1^0}(u_1^0),$$

and thus,

$$P\{y_1 = k, y_2 = 1, y_3 = 0\} = \frac{1}{\sqrt{2\pi}\sigma_1^0}\frac{1}{k!}\mu^{0k}\int_{-\infty}^{+\infty}\left\{\left[1-\Phi\left(\frac{-x_2'\beta_2^0-\alpha_{21}^0 k-\frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}\right)\right]\right.$$
$$\left[\Phi\left(\frac{-x_3'\beta_3^0-\alpha_{31}^0 k-\alpha_{32}^0-\frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right. \tag{2.68}$$
$$\left.\left.-\Psi\left(\frac{\rho_{23}^0-\rho_{12}^0\rho_{13}^0}{\sqrt{(1-\rho_{12}^{02})(1-\rho_{13}^{02})}}, \frac{-x_2'\beta_2^0-\alpha_{21}^0 k-\frac{\rho_{12}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{12}^{02}}}, \frac{-x_3'\beta_3^0-\alpha_{31}^0 k-\alpha_{32}^0-\frac{\rho_{13}^0}{\sigma_1^0}u_1^0}{\sqrt{1-\rho_{13}^{02}}}\right)\right]\right\} \tag{2.69}$$
$$\times \exp\left[-\left(\mu^0 \exp u_1^0\right) + k u_1^0\right] \exp\left(-\frac{u_1^{02}}{2\sigma_1^{02}}\right) du_1^0 \tag{2.70}$$

The log-likelihood function. The log-likelihood function for a sample with N observation may be written as

$$\mathcal{L}_N = \sum_{n=1}^{N}\sum_{k=0}^{\infty}\sum_{i=0}^{1}\sum_{j=1}^{5} z_{nkij}\log P_{nkij} \tag{2.71}$$

where

$$z_{nkij} = \begin{cases} 1 & \text{if } y_{n1} = k,\ y_{n2} = i \text{ and } y_{n3} = j \quad (k = 0, 1, \ldots, i = 0, 1 \text{ and } j = 0, 1) \\ 0 & \text{otherwise} \end{cases} \tag{2.72}$$

and P_{nij} is defined as

$$P_{nkij} = P\{z_{nkij} = 1\} \tag{2.73}$$

that is to say

$$\begin{aligned} P_{nkij} &= P(y_{n1} = k, y_{n2} = i, y_{n3} = j) \\ &= \int_{-\infty}^{+\infty} P(y_{n2} = i, y_{n3} = j \mid y_{n1} = k, u_{n1}) P(y_{n1} = k \mid u_{n1}) d\varphi_{\sigma_1}(u_{n1}) \end{aligned} \tag{2.74}$$

where $d\varphi_{\sigma_1^0}(u_1^0)$ is defined as

$$d\varphi_{\sigma_1}(u_{n1}) = \frac{1}{\sigma_1}\varphi\left(\frac{u_{n1}}{\sigma_1}\right) du_{n1}.$$

Explicitly,

$$P_{nkij} = \frac{1}{\sqrt{2\pi}\sigma_1}\int_{-\infty}^{+\infty} P(y_{n2} = i, y_{n3} = j \mid y_{n1} = k, u_{n1}) P(y_{n1} = k \mid u_{n1}) \exp\left(-\frac{u_{n1}^2}{2\sigma_1^2}\right) du_{n1}. \tag{2.75}$$

From a numerical point of view, it is convenient to make the change of variable $v_n = u_{n1}/\sqrt{2}\sigma_1$ so that (2.75) may be written as

$$P_{nkij} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} P(y_{n2} = i, y_{n3} = j \mid y_{n1} = k, \sqrt{2}\sigma_1 v_n) P(y_{n1} = k \mid \sqrt{2}\sigma_1 v_n) \exp\left(-v_n^2\right) dv_n. \tag{2.76}$$

References

[1] Greene, William (1997), "FIML estimation of sample selection models for count data", Stern School of Business, New York University. *Manuscript*

[2] Huguenin, Jacques (2004), *Multivariate normal distribution and simultaneous equation Probit analysis*, PhD Thesis, Department of Economics (DEEP), University of Lausanne, Switzerland.

[3] Huguenin, Jacques, Florian Pelgrin and Alberto Holly (2013), "Estimation of Multivariate Probit Models by Exact Maximum Likelihood", Working Paper

[4] Lazard Holly, Hippolyte and Alberto Holly (2002), "Computation of the probability that a d–dimensional normal vector belongs to a polyhedral cone with arbitrary vertex", *Mimeo*

[5] Miranda, Alfonso (2004), "FIML estimation of an endogenous switching model for count data", *The Stata Journal*, 4, Number 1, pp. 40–49