Appendix: "Treatment Effects and Informative Missingness with an Application to Bank Recapitalization Programs"

Angela Vossmeyer

University of California, Irvine*

This appendix contains detailed instructions for the Markov chain Monte Carlo (MCMC) algorithm employed in the paper. The appendix first describes the steps for the algorithm followed by definitions of the vectors and matrices involved.

MCMC Estimation Algorithm for Censored Outcomes

1. Sample $\boldsymbol{\beta}$ from the distribution $\boldsymbol{\beta}\,|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\theta} \setminus \boldsymbol{\beta}$.[1]
2. Sample $\boldsymbol{\Omega}$ from the distribution $\boldsymbol{\Omega}|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\theta} \setminus \boldsymbol{\Omega}$ in a 1-block, multi-step procedure.
3. For $i \in N_1,$ sample $\mathbf{y}_1^*$ from the distribution $\mathbf{y}_1^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_1^*$.
4. For $i \in N_2,$ sample $\mathbf{y}_2^*$ from the distribution $\mathbf{y}_2^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_2^*$.
5. For $i : y_{i3} = 0$, sample $\mathbf{y}_3^*$ from the distribution $\mathbf{y}_3^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_3^*$.
6. For $i : y_{i4} = 0$, sample $\mathbf{y}_4^*$ from the distribution $\mathbf{y}_4^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_4^*$.
7. For $i : y_{i5} = 0$, sample $\mathbf{y}_5^*$ from the distribution $\mathbf{y}_5^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_5^*$.

*Department of Economics, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100; e-mail: angela.vossmeyer@uci.edu.

[1]The notation "\" represents "except", e.g., $\mathbf{y}^* \setminus \mathbf{y}_1^*$ says all elements in $\mathbf{y}^*$ except $\mathbf{y}_1^*$.

Step 1: Sampling $\boldsymbol{\beta}$
Sample $\boldsymbol{\beta}\,|\mathbf{y}^*, \boldsymbol{\theta} \setminus \boldsymbol{\beta} \sim \mathcal{N}(\mathbf{b}, \mathbf{B})$, where

$$\begin{aligned}
\mathbf{b} &= \mathbf{B}(\mathbf{B}_0^{-1}\mathbf{b}_0 + \sum_{i \in N_1} \mathbf{J}_C'\mathbf{X}_{iC}'\boldsymbol{\Omega}_C^{-1}\mathbf{y}_{iC}^* + \\
&\quad \sum_{i \in N_2} \mathbf{J}_D'\mathbf{X}_{iD}'\boldsymbol{\Omega}_D^{-1}\mathbf{y}_{iD}^* + \sum_{i \in N_3} \mathbf{J}_A'\mathbf{X}_{iA}'\boldsymbol{\Omega}_A^{-1}\mathbf{y}_{iA}^*), \\
\mathbf{B} &= (\mathbf{B}_0^{-1} + \sum_{i \in N_1} \mathbf{J}_C'\mathbf{X}_{iC}'\boldsymbol{\Omega}_C^{-1}\mathbf{X}_{iC}\mathbf{J}_C + \\
&\quad \sum_{i \in N_2} \mathbf{J}_D'\mathbf{X}_{iD}'\boldsymbol{\Omega}_D^{-1}\mathbf{X}_{iD}\mathbf{J}_D + \sum_{i \in N_3} \mathbf{J}_A'\mathbf{X}_{iA}'\boldsymbol{\Omega}_A^{-1}\mathbf{X}_{iA}\mathbf{J}_A)^{-1}.
\end{aligned}$$

Step 2: Sampling $\boldsymbol{\Omega}$
Sample $\boldsymbol{\Omega}|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\theta} \setminus \boldsymbol{\Omega}$ in a 1-block, nine-step procedure by drawing Ω_{11} , $\boldsymbol{\Omega}_{tt\cdot l} = \boldsymbol{\Omega}_{tt} - \boldsymbol{\Omega}_{tl}\boldsymbol{\Omega}_{ll}^{-1}\boldsymbol{\Omega}_{lt}$, and $\mathbf{B}_{lt} = \boldsymbol{\Omega}_{ll}^{-1}\boldsymbol{\Omega}_{lt}$, and then reconstructing $\boldsymbol{\Omega}$ from these quantities

2. (a) $\Omega_{11}|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\theta} \setminus \boldsymbol{\Omega} \sim \mathcal{IW}(\nu - 1 + n,\ Q_{11} + \sum_{N_1,N_2,N_3} \boldsymbol{\eta}_{i1}^* \boldsymbol{\eta}_{i1}^{*'})$
   i. $\boldsymbol{\eta}_{i1}^* = y_{i1}^* - \mathbf{x}_{i1}\mathbf{J}_1\boldsymbol{\beta}$, where $\mathbf{J}_1 = \begin{bmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \end{bmatrix}_{k_1 \times K}$
   (b) $\Omega_{22\cdot 1}|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\theta} \setminus \boldsymbol{\Omega} \sim \mathcal{IW}(\nu + n_2 + n_3, R_{22\cdot 1})$
   (c) $B_{12}|\mathbf{y}, \mathbf{y}^*, \Omega_{22\cdot 1} \sim \mathcal{MN}(R_{11}^{-1}R_{21}, \Omega_{22\cdot 1} \otimes R_{11}^{-1})$
   (d) Define $\boldsymbol{\Omega}_u = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$
   (e) $\Omega_{55\cdot 1}|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\theta} \setminus \boldsymbol{\Omega} \sim \mathcal{IW}(\nu + n_1, R_{55\cdot 1})$
   (f) $B_{15}|\mathbf{y}, \mathbf{y}^*, \Omega_{55\cdot 1} \sim \mathcal{MN}(R_{11}^{-1}R_{51}, \Omega_{55\cdot 1} \otimes R_{11}^{-1})$
   (g) $\boldsymbol{\Omega}_{33\cdot u}|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\theta} \setminus \boldsymbol{\Omega} \sim \mathcal{IW}(\nu + n_2, \mathbf{R}_{33\cdot u})$
   (h) $\mathbf{B}_{u3}|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\Omega}_{33\cdot u} \sim \mathcal{MN}(\mathbf{R}_u^{-1}\mathbf{R}_{3u}, \boldsymbol{\Omega}_{33\cdot u} \otimes \mathbf{R}_u^{-1})$
   (i) $\boldsymbol{\Omega}_{44\cdot u}|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\theta} \setminus \boldsymbol{\Omega} \sim \mathcal{IW}(\nu + n_3, \mathbf{R}_{44\cdot u})$
   (j) $\mathbf{B}_{u4}|\mathbf{y}, \mathbf{y}^*, \boldsymbol{\Omega}_{44\cdot u} \sim \mathcal{MN}(\mathbf{R}_u^{-1}\mathbf{R}_{4u}, \boldsymbol{\Omega}_{44\cdot u} \otimes \mathbf{R}_u^{-1})$

where $\mathbf{R} = \mathbf{Q} + \sum \boldsymbol{\eta}_i^* \boldsymbol{\eta}_i^{*\prime}$, and the following subsections are obtained by partitioning $\mathbf{R}$ to conform to $\mathbf{Q}$, and $\mathbf{R}_{tt\cdot l} = \mathbf{R}_{tt} - \mathbf{R}_{tl}\mathbf{R}_{ll}^{-1}\mathbf{R}_{lt}$. From these sampling densities, $\boldsymbol{\Omega}$ can be recovered.

Steps 3-7: Sampling $\mathbf{y}^*$
For censored outcomes, $\mathbf{y}^*$ is sampled following from a truncated normal,

$$\begin{aligned}
\mathbf{y}_1^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_1^* &\sim \mathcal{TN}_{(-\infty,0)}(\mathbf{x}_{i1}'\boldsymbol{\beta}_1 + E(\varepsilon_{i1}|\varepsilon_{i\setminus 1}), var(\varepsilon_{i1}|\varepsilon_{i\setminus 1})), \quad i \in N_1, \\
\mathbf{y}_2^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_2^* &\sim \mathcal{TN}_{(-\infty,0)}(\mathbf{x}_{i2}'\boldsymbol{\beta}_2 + E(\varepsilon_{i2}|\varepsilon_{i\setminus 2}), var(\varepsilon_{i2}|\varepsilon_{i\setminus 2})), \quad i \in N_2, \\
\mathbf{y}_3^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_3^* &\sim \mathcal{TN}_{(-\infty,0)}((\mathbf{x}_{i3}'\ y_{i1})\boldsymbol{\beta}_3 + E(\varepsilon_{i3}|\varepsilon_{i\setminus 3}), var(\varepsilon_{i3}|\varepsilon_{i\setminus 3})), \quad i : y_{i3} = 0, \\
\mathbf{y}_4^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_4^* &\sim \mathcal{TN}_{(-\infty,0)}((\mathbf{x}_{i4}'\ y_{i1}\ y_{i2})\boldsymbol{\beta}_4 + E(\varepsilon_{i4}|\varepsilon_{i\setminus 4}), var(\varepsilon_{i4}|\varepsilon_{i\setminus 4})), \quad i : y_{i4} = 0, \\
\mathbf{y}_5^*|\mathbf{y}, \boldsymbol{\theta}, \mathbf{y}^* \setminus \mathbf{y}_5^* &\sim \mathcal{TN}_{(-\infty,0)}(\mathbf{x}_{i5}'\boldsymbol{\beta}_5 + E(\varepsilon_{i5}|\varepsilon_{i\setminus 5}), var(\varepsilon_{i5}|\varepsilon_{i\setminus 5})), \quad i : y_{i5} = 0.
\end{aligned}$$

Definitions

Priors: It is assumed that $\boldsymbol{\beta}$ has a joint normal distribution with mean $\boldsymbol{\beta}_0$ and variance $\mathbf{B}_0$ and (independently) that the covariance matrix $\boldsymbol{\Omega}$ has an inverse Wishart distribution with parameters υ and $\mathbf{Q}$,

$$\pi(\boldsymbol{\beta}, \boldsymbol{\Omega}) = \mathcal{N}(\boldsymbol{\beta} \,|\boldsymbol{\beta}_0, \,\mathbf{B}_0)\mathcal{IW}(\boldsymbol{\Omega}|\upsilon, \mathbf{Q}).$$

Data: For the i-th observation, define the following vectors and matrices,

$$\mathbf{y}^*_{iC} = (y^*_{i1}, y^*_{i5})', \quad \mathbf{y}^*_{iD} = (y^*_{i1}, y^*_{i2}, y^*_{i3})', \quad \mathbf{y}^*_{iA} = (y^*_{i1}, y^*_{i2}, y^*_{i4})',$$

$$\mathbf{X}_{iC} = \begin{pmatrix} \mathbf{x}'_{i1} & 0 \\ 0 & \mathbf{x}'_{i5} \end{pmatrix}, \quad \mathbf{X}_{iD} = \begin{pmatrix} \mathbf{x}'_{i1} & 0 & 0 \\ 0 & \mathbf{x}'_{i2} & 0 \\ 0 & 0 & (\mathbf{x}'_{i3}\ y_{i1}) \end{pmatrix}, \quad \mathbf{X}_{iA} = \begin{pmatrix} \mathbf{x}'_{i1} & 0 & 0 \\ 0 & \mathbf{x}'_{i2} & 0 \\ 0 & 0 & (\mathbf{x}'_{i4}\ y_{i1}\ y_{i2}) \end{pmatrix}.$$

Let $N_1 = \{i : y_{i1} = 0\}$ be the n_1 observations in the non-selected sample, and $N_2 = \{i : y_{i1} > 0 \ \textit{and} \ y_{i2} = 0\}$ be the n_2 observations in the selected untreated sample. Set $N_3 = \{i : y_{i1} > 0 \ \textit{and} \ y_{i2} > 0\}$ to be the n_3 observations in the selected treated sample. Let $\boldsymbol{\theta}$ be the set of all model parameters.

In order to isolate the observed vectors and matrices that correspond to the 3 different subsets of the sample, define

$$\mathbf{J}_C = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I} \end{pmatrix}_{(k_1+k_5)\times K}, \ \mathbf{J}_D = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 & 0 \end{pmatrix}_{(k_1+k_2+k_3)\times K},$$
$$\mathbf{J}_A = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{I} & 0 \end{pmatrix}_{(k_1+k_2+k_4)\times K}$$

where $K = k_1 + k_2 + k_3 + k_4 + k_5$, which represents the number of covariates in each equation.

For $i \in N_1$ (non-selected sample),

$$\boldsymbol{\eta}^*_{iC} \;=\; \mathbf{y}^*_{iC} - \mathbf{X}_{iC}\mathbf{J}_C\boldsymbol{\beta},$$

for $i \in N_2$ (selected untreated sample),

$$\boldsymbol{\eta}^*_{iD} \;=\; \mathbf{y}^*_{iD} - \mathbf{X}_{iD}\mathbf{J}_D\boldsymbol{\beta},$$

and for $i \in N_3$ (selected treated sample),

$$\boldsymbol{\eta}^*_{iA} \;=\; \mathbf{y}^*_{iA} - \mathbf{X}_{iA}\mathbf{J}_A\boldsymbol{\beta},$$