Introduction to Bayesian Econometrics, 2nd Edition
Answers to Exercises, Chapter 12

1. Since y_{ji} and s_i^* have a joint normal distribution, we can use the result from Appendix A that $f(y_{ji}|s_i^*)$ has a normal distribution with mean

$$\begin{aligned} E(y_{ji}|s_i^*) &= E(y_{ji}) + \omega_j\big(E(s_i^*) - s_i^*\big) \\ &= x_i'\beta_j + \omega_j\big(x_i'\gamma_1 + z_i'\gamma_2 - s_i^*\big) \\ &= x_i'\beta_j + \omega_j v_i, \end{aligned}$$

and variance

$$\begin{aligned} \mathrm{Var}(y_{ji}|s_i^*) &= \sigma_{jj} + \omega_j^2 - \omega_j^2 \\ &= \sigma_{jj}. \end{aligned}$$

2. As usual, we find the posterior distribution of $\phi_j = (\beta_j, \omega_j)$ by multiplying the likelihood times the prior. Let θ contain all the parameters. Then the likelihood

$$\prod_i f(y_{ij}, s_i^*|\theta) = \prod_i f(y_{ij}|s_i^*, \theta) f(s_i^*|\theta),$$

but s_i^* depends on neither β_j nor ω_j. Therefore

$$\pi(\beta_j, \omega_j|Y_j, \gamma, S_j^*, \sigma_{jj}) \propto f(Y_j|S_j^*, \beta_j, \omega_j, \sigma_{jj})\, \pi(\beta_j)\, \pi(\omega_j).$$

From here, we have

$$\begin{aligned}\pi(\beta_j, \omega_j | Y_j, \gamma, S_j^*, \sigma_{jj}) &\propto \exp\left[-\frac{1}{2\sigma_{jj}}(Y_j - X_j\beta_j - V_j\omega_j)'(Y_j - X_j\beta_j - V_j\omega_j)\right] \\ &\quad \times \exp\left[-\frac{1}{2}(\beta_j - b_{j0})' B_{j0}^{-1}(\beta_j - b_{j0})\right] \\ &\quad \times \exp\left[-\frac{1}{2}(\omega_j - m_{j0})' M_{j0}^{-1}(\omega_j - m_{j0})\right] \\ &= \exp\left[-\frac{1}{2\sigma_{jj}}(Y_j - W_j\phi_j)'(Y_j - W_j\phi_j)\right] \\ &\quad \times \exp\left[-\frac{1}{2}(\phi_j - f_{j0})' F_{j0}^{-1}(\phi_j - f_{j0})\right].\end{aligned}$$

The usual calculations show that the full conditional distribution of ϕ_j is Gaussian with covariance matrix

$$\mathrm{Cov}(\phi_j | Y_j, \gamma, S_j^*, \sigma_{jj}) = \left[\sigma_{jj}^{-1} W_j' W_j + F_{j0}^{-1}\right]^{-1}$$

and mean

$$E(\phi_j | Y_j, \gamma, S_j^*, \sigma_{jj}) = \mathrm{Cov}(\phi_j | Y_j, \gamma, S_j^*, \sigma_{jj}) \left[\sigma_{jj}^{-1} W_j' Y_j + F_{j0}^{-1} f_{j0}\right].$$

3. We can use the results of the previous exercise to write

$$\begin{aligned}\pi(\sigma_{jj} | Y_j, S_j^*, \beta_j, \gamma, \omega_j) &\propto f(Y_j | S_j^*, \beta_j, \gamma, \omega_j, \sigma_{jj})\, \pi(\sigma_{jj}) \\ &\propto \left(\frac{1}{\sigma_{jj}}\right)^{n_j/2} \exp\left[-\frac{1}{2\sigma_{jj}}(Y_j - W_j\phi_j)'(Y_j - W_j\phi_j)\right] \\ &\quad \times \left(\frac{1}{\sigma_{jj}}\right)^{\nu_{j0}/2+1} \exp\left[-\frac{d_{j0}}{2\sigma_{jj}}\right],\end{aligned}$$

which implies that σ_{jj} has an inverse gamma distribution with parameters

$$\frac{n_j + \nu_{j0}}{2}, \quad \frac{d_{j0} + (Y_j - W_j\phi_j)'(Y_j - W_j\phi_j)}{2}.$$

4. From the conditional distribution of a joint Gaussian distribution, we have that $s_i^*|y_{ji}$ has a truncated Gaussian distribution with expected value

$$\begin{aligned}
E(s_i^*|y_{ji}) &= E(s_i^*) + \frac{\omega_j}{\sigma_{jj}+\omega_j^2}\big(y_{ji} - E(y_{ji})\big) \\
&= x_i'\gamma_1 + z_i'\gamma_2 + \frac{\omega_j}{\sigma_{jj}+\omega_j^2}\, u_{ji},
\end{aligned}$$

and variance

$$1 - \frac{\omega_j^2}{\sigma_{jj}+\omega_j^2} = \frac{\sigma_{jj}}{\sigma_{jj}+\omega_j^2}.$$

The distribution is truncated over $(-\infty, 0]$ for $s_i = 0$ and $(0, \infty)$ for $s_i = 1$.

5. From the definitions in the text, we have, for $j = 1, 2$,

$$S_j^* = P_j\gamma + V_j,$$

where V_j is the vector of v_{ji} for $i \in N_j$. The likelihood function for $S_j^*|Y_j$ is therefore Gaussian with mean

$$\begin{aligned}
E(S_j^*|Y_j) &= E(S_j^*) + \frac{\omega_j}{\sigma_{jj}+\omega_j^2}\big(Y_j - E(Y_j)\big) \\
&= P_j\gamma + \frac{\omega_j}{\sigma_{jj}+\omega_j^2} U_j,
\end{aligned}$$

and variance

$$1 - \frac{\omega_j^2}{\sigma_{jj}+\omega_j^2} = \frac{\sigma_{jj}}{\sigma_{jj}+\omega_j^2}.$$

It follows that

$$\begin{aligned}
f(S_j^*|Y_j,\theta) &\propto \exp\left[-\frac{\sigma_{jj}+\omega_j^2}{2\sigma_{jj}}\left(S_j^* - P_j\gamma - \frac{\omega_{jj}}{\sigma_{jj}+\omega_j^2}\, U_j\right)'\left(S_j^* - P_j\gamma - \frac{\omega_{jj}}{\sigma_{jj}+\omega_j^2}\, U_j\right)\right], \\
f(T_j^*|Y_j,\theta) &\propto \exp\left[-\frac{\sigma_{jj}+\omega_j^2}{2\sigma_{jj}}\left(T_j^* - P_j\gamma\right)'\left(T_j^* - P_j\gamma\right)\right],
\end{aligned}$$

where θ contains the relevant parameters. Letting $j = 1, 2$ and introducing the prior distribution for γ, we have

$$\begin{aligned}\pi(\gamma|y, s^*, \beta_0, \beta_1, \sigma_{00}, \sigma_{11}, \omega_{00}, \omega_{11}) \propto \exp&\left[-\frac{\sigma_{00} + \omega_0^2}{2\sigma_{00}} \left(T_0^* - P_0\gamma\right)' \left(T_0^* - P_0\gamma\right)\right] \\ \times \exp&\left[-\frac{\sigma_{11} + \omega_1^2}{2\sigma_{11}} \left(T_1^* - P_1\gamma\right)' \left(T_1^* - P_1\gamma\right)\right] \\ \times \exp&\left[-\frac{1}{2}(\gamma - g_0)'G_0^{-1}(\gamma - g_0)\right].\end{aligned}$$

To get a compact representation of this posterior distribution, define

$$a_j = \frac{\sigma_{jj} + \omega_j^2}{\sigma_{jj}}, \quad j = 1, 2.$$

We can then write

$$\begin{aligned}&a_0(T_0^* - P_0\gamma)'(T_0^* - P_0\gamma) + a_1(T_1^* - P_1\gamma)'(T_1 * -P_1\gamma) \\ &\qquad = \gamma'(a_0P_0'P_0 + a_1P_1'P_1)\gamma - 2\gamma'(a_0P_0'T_0^* + a_1P_1'T_1^*) + C \\ &\qquad \equiv \gamma'P\gamma - 2\gamma'T^* + C,\end{aligned}$$

where C does not depend on γ. This allows us to combine the first two terms in the posterior to rewrite

$$\begin{aligned}&\pi(\gamma|y, s^*, \beta_0, \beta_1, \sigma_{00}, \sigma_{11}, \omega_0, \omega_1) \\ &\qquad \propto \exp\left[-\frac{1}{2}(\gamma'P\gamma - 2\gamma'T^* + C)\right] \exp\left[-\frac{1}{2}(\gamma - g_0)'G_0^{-1}(\gamma - g_0)\right],\end{aligned}$$

from which we conclude that the full conditional distribution of γ, given the data and other paramaters is Gaussian, with covariance matrix

$$\left[P + G_0^{-1}\right]^{-1}$$

and mean vector

$$\left[P + G_0^{-1}\right]^{-1} \left(T^* + G_0^{-1} g_0\right).$$

6. The Gibbs algorithm can be constructed from the full conditional distributions derived in the previous four exercises.

7. For $j = 1$, we have

$$\begin{aligned}
f(y_{1i}, s_i = 1|\beta_1^*, \gamma^*, \sigma_{11}^*, \omega_1^*) &= f(y_{1i}|\beta_1^*, \sigma_{11}^*, \omega_1^*)\, P(s_i = 1|y_{1i}, \beta_1^*, \gamma^*, \sigma_{11}^*, \omega_1^*) \\
&= f(y_{1i}|\beta_1^*, \sigma_{11}^*, \omega_1^*) \int_0^\infty f(s_i^*|\mu_{1i}^*, \psi_1^*)\, ds_i^* \\
&= N(y_{1i}|x_i'\beta_1^*, \sigma_{11}^* + \omega_1^{*2}) \left(1 - \Phi\left(-\frac{\mu_{1i}^*}{\psi_1^*}\right)\right) \\
&= N(y_{1i}|x_i'\beta_1^*, \sigma_{11}^* + \omega_1^{*2})\, \Phi\left(\frac{\mu_{1i}^*}{\psi_1^*}\right),
\end{aligned}$$

which verifies (12.9).

8. If the response variable is binary, we can write, for $j = 0, 1$,

$$\begin{aligned}
y_{ij}^* &= x_i'\beta_j + u_{ji}, \\
y_{ij} &= \begin{cases} 0, & \text{if } y_{ij}^* \leq 0, \\ 1, & \text{if } y_{ij}^* > 0, \end{cases} \\
y_i^* &= (1 - s_i) y_{0i}^* + s_i y_{1i}.
\end{aligned}$$

In computing the ATE, $E(y_{ji}|s_i)$ is regarded as $P(y_{ij}|s_i) = 1$.

For identification purposes, $\mathrm{Var}(y_{ij}^*) = 1$, so that

$$\mathrm{Cov}(u_{ji}, v_i) \equiv \Sigma_j = \begin{pmatrix} 1 & \omega_j \\ \omega_j & 1 \end{pmatrix},$$

so that it is not necessary to sample for the σ_{jj}. Moreover, $\omega_j^2 < 1$.

We can now write

$$
\begin{aligned}
f(y_i, s_i | s_i^*, y_i^*, \theta) &= f(y_i | s_i^*, y_i^*, \theta)\, f(s_i | s_i^*, \theta) \\
&= [1(y_i = 0 | s_i) 1(y_i^* \leq 0 | s_i) + 1(y_i = 1 | s_i) 1(y_i^* > 0 | s_i)] \\
&\quad \times [1(s_i = 0) 1(s_i^* \leq 0) + 1(s_i = 1) 1(s_i^* > 0)].
\end{aligned}
$$

The full conditional distributions needed to devise a Gibbs sampler are those in the text, with the following changes: no sampling is done for the σ_{jj}; Y_j^* replaces Y_j in the full conditional distributions of (β_j, ω_j), s_i^*, and γ; and steps are added for the sampling of y_{0i}^* and y_{1i}^* from the usual truncated Gaussian distributions.

9. Set

$$
\begin{aligned}
y_i^* &= x_i'\beta_1 + \beta_s x_{is} + u_i, \\
y_i &= \begin{cases} 0, & \text{if } y_i^* \leq 0, \\ 1, & \text{if } y_i^* > 0, \end{cases}
\end{aligned}
$$

and $\sigma_{11} = 1$. As before, we include instrumental variables z_i, and specify

$$
x_{is} = x_i'\gamma_1 + z_i'\gamma_2 + v_i.
$$

To avoid having to deal with constrained estimators, we reparamaterize the covariance matrix as in section 12.1,

$$
\Sigma = \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_{22} + \sigma_{12}^2 \end{pmatrix}.
$$

Notice that σ_{12} is unconstrained and that $\sigma_{22} > 0$.

We can now proceed as follows. First, write the likelihood function in terms of

y^* as

$$f(y_i^*, x_{is}|\theta) = f(x_{is}|y_i^*, \theta)\, f(y_i^*|\theta).$$

The first expression can be written as

$$f(x_{is}|y_i^*, \theta) \propto \exp\left[-\frac{1}{2\sigma_{22}}\left(x_{is} - x_i'\gamma_1 - z_i'\gamma_2 - \sigma_{12}\hat{u}_i\right)'\left(x_{is} - x_i'\gamma_1 - z_i'\gamma_2 - \sigma_{12}\hat{u}_i\right)\right],$$

where $\hat{u}_i = y_i^* - x_i'\beta_1 - \beta_s x_{is}$. We can now find the full conditional distribution of $\gamma = (\gamma_1, \gamma_2, \sigma_{12})$ by writing

$$\begin{aligned}\pi(\gamma|y^*, \beta_1, \beta_s, \sigma_{22}) \propto \exp&\left[-\frac{1}{2\sigma_{22}}(x_s - W\gamma)'(x_s - W\gamma)\right] \\ &\times \exp\left[-(\gamma - g_0)'G_0^{-1}(\gamma - g_0)\right],\end{aligned}$$

where γ has been given a joint Gaussian prior with mean g_0 and covariance G_0 and W is an $n \times (K_1 + K_2 + 1)$ matrix with ith row $(x_i', z_i', \hat{u}_i)$. The full conditional for γ is now easily derived.

For σ_{22}^{-1}, we take a gamma prior with parameters $\alpha_{20}/2, \delta_{20}/2$:

$$\begin{aligned}\pi(\sigma_{22}|y^*, \gamma, \beta_1, \beta_s) \propto &\left(\frac{1}{2\sigma_{22}}\right)^{n/2} \exp\left[-\frac{1}{2\sigma_{22}}(x_s - W\gamma)'(x_s - W\gamma)\right] \\ &\times \left(\frac{1}{\sigma_{22}}\right)^{\alpha_0/2+1} \exp\left[-\frac{\delta_0}{2\sigma_{22}}\right].\end{aligned}$$

The full conditional can be found by standard methods as a gamma distribution.

For $\beta = (\beta_1, \beta_s)$, we write the likelihood function as

$$f(y_i^*, x_{is}|\theta) = f(y_i|x_{is}, \theta)\, f(x_{is}\theta).$$

Given x_{is}, we can compute $\hat{v}_i = x_{is} - x_i'\gamma_1 - z_i'\gamma_2$. Then, with $X = (x, x_s)$ and

$$\omega_{12} = \frac{\sigma_{22}}{\sigma_{22} + \sigma_{12}^2},$$

$$\pi(\beta | y^*, x_s, \gamma, \sigma_{22}) \propto \exp\left[-\frac{\sigma_{22} + \sigma_{12}^2}{2\sigma_{12}^2}\big(y^* - X\beta - \omega_{12}\hat{v}\big)'\big(y^* - X\beta - \omega_{12}\hat{v}\big)\right]$$
$$\times \exp\left[-\frac{1}{2}(\beta - b_0)'B_0^{-1}(\beta - b_0)\right].$$

From here, it is clear that β is distributed as Gaussian, with parameters that can be derived in the usual way.

Finally, we include y_i^* in the sampler. Given the relevant parameters, it has a truncated normal distribution that depends on y_i in the usual way, with

$$E(y_i^* | \theta) = x_i'\beta_1 + \beta_s x_{is}$$

and a variance of 1.

10. We write the version with a Tobit selection model as

$$\begin{aligned} y_i &= x_i'\beta_1 + u_i, \\ s_i^* &= x_i'\gamma_1 + z_i'\gamma_2 + v_i, \\ s_i &= \begin{cases} 0, & \text{if } s_i^* \le 0, \\ s_i^*, & \text{if } s_i^* > 0, \end{cases} \end{aligned}$$

where $(u_i, v_i) \sim N_2(0, \Sigma)$, and

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}.$$

This problem is studied in a more general form in Chib, Greenberg, and Jeliazkov (2009). The main issue is that of estimating Σ, which is now defined as

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}.$$

Since Σ is required to be positive definite, we must operate with the whole matrix. It is, however, possible to work with three parameters, from which the full matrix may be recovered; they are

$$\begin{aligned} \sigma_{11\cdot 2} &= \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}, \\ \beta_{12} &= \frac{\sigma_{12}}{\sigma_{22}}, \end{aligned}$$

and σ_{22}. Assume for now that the elements of Σ are available.

First, we add s_i^* to the sampler for the observations for which $s_i = 0$. For them, $s_i^* \sim \mathrm{TN}(x_i'\gamma_1 + z_i'\gamma_2, \sigma_{22})$, truncated to $(-\infty, 0]$.

For β, we include σ_{22}, is not required to equal one:

$$\begin{aligned} \pi(\beta|y, s^*, \Sigma) \propto \exp &\left[-\frac{1}{2\sigma_{22}} \sum_{i \in N_0} (\eta_i - X_i\beta)' J' J (\eta_i - X_i\beta) \right] \\ &\times \exp \left[-\frac{1}{2} \sum_{i \in N_1} (\eta_i - X_i\beta)' \Sigma^{-1} (\eta_i - X_i\beta) \right], \end{aligned}$$

so that

$$\beta | y, s^*, \Sigma \sim N_{K_1+K_2}(\bar{\beta}, B_1),$$

where

$$B_1 = \left[\sigma_{22}^{-1} \sum_{i \in N_0} X_i' J X_i + \sum_{i \in N_1} X_i' \Sigma^{-1} X_i + B_0^{-1} \right]^{-1},$$

$$\bar{\beta} = B_1 \left[\sigma_{22}^{-1} \sum_{i \in N_0} X_i' J \eta_i + \sum_{i \in N_1} X_i' \Sigma^{-1} \eta_i + B_0^{-1} \beta_0 \right].$$

The sampling for Σ is discussed in detail in the article mentioned above; the sampling uses equation (A.14). It is accomplished by assuming an inverse Wishart distribution, and then showing that σ_{22} and $\sigma_{11 \cdot 2}$ have inverse Wishart distributions and β_{12} has a normal distribution, with parameters specified in the article.