Problem 1

Part i)

Our model is given by: $Y = X\beta + e$, and $p\left(e \mid \sigma^2\right) \sim \mathcal{N}\left(0, \sigma^2 I_n\right)$. Assuming that X is exogenous, the likelihood is, up to a constant, given by:

$$p\left(y \mid \beta, h\right) \propto h^{\frac{n}{2}} \exp\left(-\frac{h}{2}\left(Y - X\beta\right)'\left(Y - X\beta\right)\right)$$

Our prior for β and h follows a Normal-Gamma distribution with parameters: $\left(\underline{\beta}, \underline{V}, \underline{h}, \underline{\nu}\right)$, that is:

$$p\left(\beta, h\right) \propto h^{\frac{\nu-2}{2}} \exp\left(\frac{-\underline{\nu} h}{2\underline{h}}\right) \exp\left(-\frac{h}{2}\left(\beta - \underline{\beta}\right)' \underline{V}^{-1}\left(\beta - \underline{\beta}\right)\right)$$

Using Bayes' Rule:

$$\begin{aligned} p\left(\beta, h \mid y\right) &\propto p\left(y \mid \beta, h\right) p\left(\beta, h\right) \\ &= h^{\frac{n}{2}} \exp\left(-\frac{h}{2}\left(Y - X\beta\right)'\left(Y - X\beta\right)\right) h^{\frac{\nu-2}{2}} \exp\left(\frac{-\underline{\nu} h}{2\underline{h}}\right) \exp\left(-\frac{h}{2}\left(\beta - \underline{\beta}\right)' \underline{V}^{-1}\left(\beta - \underline{\beta}\right)\right) \end{aligned}$$

Part ii)

Again by Bayes' Rule:

$$p\left(\beta \mid h, y\right) = \frac{p\left(y \mid \beta, h\right) p\left(\beta, h\right)}{p\left(h, y\right)}$$

Since we are interest in β we can put $p\left(h\right)$ in the constant term and write:

$$\begin{aligned} p\left(\beta \mid h, y\right) &\propto p\left(y \mid \beta, h\right) p\left(\beta, h\right) \\ &= h^{\frac{n}{2}} \exp\left(-\frac{h}{2}\left(Y - X\beta\right)'\left(Y - X\beta\right)\right) h^{\frac{\nu-2}{2}} \exp\left(\frac{-\underline{\nu} h}{2\underline{h}}\right) \exp\left(-\frac{h}{2}\left(\beta - \underline{\beta}\right)' \underline{V}^{-1}\left(\beta - \underline{\beta}\right)\right) \\ &\propto \exp\left(-\frac{h}{2}\left(Y - X\beta\right)'\left(Y - X\beta\right)\right) \exp\left(-\frac{h}{2}\left(\beta - \underline{\beta}\right)' \underline{V}^{-1}\left(\beta - \underline{\beta}\right)\right) \end{aligned}$$

We will rewrite the expression:$\left(Y - X\beta\right)'\left(Y - X\beta\right)$. Now let $\hat{\beta} = \left(X'X\right)^{-1} X'Y$ and $\hat{e} = Y - X\hat{\beta}$, we then have:

$$\begin{aligned} \left(Y - X\beta\right)'\left(Y - X\beta\right) &= \left(Y - X\hat{\beta} - X\left(\beta - \hat{\beta}\right)\right)'\left(Y - X\hat{\beta} - X\left(\beta - \hat{\beta}\right)\right) \\ &= \left(\hat{e} - X\left(\beta - \hat{\beta}\right)\right)'\left(\hat{e} - X\left(\beta - \hat{\beta}\right)\right) \\ &= \hat{e}'\hat{e} + \left(\beta - \hat{\beta}\right)' X'X\left(\beta - \hat{\beta}\right) \end{aligned}$$

Since $\hat{e}'\hat{e}$ is not a function of β, we can neglect that term and write:

$$\begin{aligned}
p(\beta \mid h, y) &\propto \exp\left(-\frac{1}{2}\left(\beta-\hat{\beta}\right)' hX'X\left(\beta-\hat{\beta}\right)\right)\exp\left(-\frac{h}{2}\left(\beta-\underline{\beta}\right)'\underline{V}^{-1}\left(\beta-\underline{\beta}\right)\right) \\
&= \exp\left(-\frac{h}{2}\left[\left(\beta-\hat{\beta}\right)' X'X\left(\beta-\hat{\beta}\right)+\left(\beta-\underline{\beta}\right)'\underline{V}^{-1}\left(\beta-\underline{\beta}\right)\right]\right) \\
&= \exp\left(-\frac{h}{2}\left[\beta'\left(X'X\right)\beta-2\beta'X'X\hat{\beta}+\hat{\beta}'X'X\hat{\beta}+\beta'\underline{V}^{-1}\beta-2\beta'\underline{V}^{-1}\underline{\beta}+\underline{\beta}'\underline{V}^{-1}\underline{\beta}\right]\right) \\
&= \exp\left(-\frac{h}{2}\left[\beta'\left(X'X+\underline{V}^{-1}\right)\beta-2\beta'\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right)+\hat{\beta}'X'X\hat{\beta}+\underline{\beta}'\underline{V}^{-1}\underline{\beta}\right]\right) \\
&\propto \exp\left(-\frac{h}{2}\left[\beta'\left(X'X+\underline{V}^{-1}\right)\beta-2\beta'\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right)\right]\right)
\end{aligned}$$

Now:

$$\begin{aligned}
\beta'\left(X'X+\underline{V}^{-1}\right)\beta-2\beta'\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right) = \beta'\left(X'X+\underline{V}^{-1}\right)\beta-2\beta'\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right)+ \\
\left[\left(X'X+\underline{V}^{-1}\right)^{-1}\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right)\right]'\left(X'X+\underline{V}^{-1}\right)^{-1}\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right) \\
-\left[\left(X'X+\underline{V}^{-1}\right)^{-1}\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right)\right]'\left(X'X+\underline{V}^{-1}\right)^{-1}\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right)
\end{aligned}$$

Note that the last term we subtract in the expression above is not a function of β so we neglect it. We then can rewrite the expression above as:

$$\left(\beta-\left(X'X+\underline{V}^{-1}\right)^{-1}\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right)\right)'\left(X'X+\underline{V}^{-1}\right)\left(\beta-\left(X'X+\underline{V}^{-1}\right)^{-1}\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right)\right)$$

We now finally define:

$$\begin{aligned}
\bar{V}^{-1} &= \left(X'X+\underline{V}^{-1}\right) \\
\bar{\beta} &= \left(X'X+\underline{V}^{-1}\right)^{-1}\left(X'X\hat{\beta}+\underline{V}^{-1}\underline{\beta}\right) \\
&= \bar{V}\left(X'y+\underline{V}^{-1}\underline{\beta}\right)
\end{aligned}$$

Then:

$$p(\beta \mid h, y) \propto \exp\left(\frac{-h}{2}\left(\beta-\bar{\beta}\right)'\bar{V}^{-1}\left(\beta-\bar{\beta}\right)\right)$$

which is the kernel of the normal density. Therefore, we can write: $\beta \mid h, y \sim \mathcal{N}\left(\bar{\beta}, h^{-1}\bar{V}\right)$

Part iii)

We will first find $p(h, \beta \mid y)$ and then integrate out β . Again, by Bayes' Rule:

$$p(h, \beta \mid y) = \frac{p(y \mid \beta, h)\, p(\beta, h)}{p(y)}$$

Since we are interest in h we can put $p(y)$ in the constant term and write:

$$\begin{aligned}
p(h,\beta \mid y) &\propto p(y \mid \beta, h)\, p(\beta, h) \\
&= h^{\frac{n}{2}} \exp\left(-\frac{h}{2}(Y-X\beta)'(Y-X\beta)\right) h^{\frac{\underline{\nu}-2}{2}} \exp\left(\frac{-\underline{\nu} h}{2\underline{h}}\right) \exp\left(-\frac{h}{2}\left(\beta-\underline{\beta}\right)' \underline{V}^{-1}\left(\beta-\underline{\beta}\right)\right) \frac{1}{(2\pi)^{\frac{k}{2}} \det\left(h^{-1}\underline{V}\right)} \\
&= h^{\frac{n}{2}} h^{\frac{\underline{\nu}-2}{2}} \exp\left(\frac{-\underline{\nu} h}{2\underline{h}}\right) \frac{1}{(2\pi)^{\frac{k}{2}} \det\left(h^{-1}\underline{V}\right)} \exp\left(-\frac{h}{2}\left[(Y-X\beta)'(Y-X\beta) + \left(\beta-\underline{\beta}\right)' \underline{V}^{-1}\left(\beta-\underline{\beta}\right)\right]\right) \\
&= h^{\frac{n}{2}} h^{\frac{\underline{\nu}-2}{2}} \exp\left(\frac{-\underline{\nu} h}{2\underline{h}}\right) \frac{1}{(2\pi)^{\frac{k}{2}} \det\left(h^{-1}\underline{V}\right)} \exp\left(-\frac{h}{2}\left[\hat{e}'\hat{e} + \left(\beta-\hat{\beta}\right)' X'X \left(\beta-\hat{\beta}\right) + \left(\beta-\underline{\beta}\right)' \underline{V}^{-1}\left(\beta-\underline{\beta}\right)\right]\right)
\end{aligned} \tag{1}$$

The same procedure we did in part ii) applies to complete the squares in $\left(\beta-\hat{\beta}\right)' X'X \left(\beta-\hat{\beta}\right) + \left(\beta-\underline{\beta}\right)' \underline{V}^{-1}\left(\beta-\underline{\beta}\right)$:

$$\begin{aligned}
\left(\beta-\hat{\beta}\right)' X'X \left(\beta-\hat{\beta}\right) + \left(\beta-\underline{\beta}\right)' \underline{V}^{-1}\left(\beta-\underline{\beta}\right) &= \beta'\left(X'X\right)\beta - 2\beta' X'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} + \beta'\underline{V}^{-1}\beta - 2\beta'\underline{V}^{-1}\underline{\beta} + \underline{\beta}'\underline{V}^{-1}\underline{\beta} \\
&= \beta'\left(X'X + \underline{V}^{-1}\right)\beta - 2\beta'\left(X'X\hat{\beta} + \underline{V}^{-1}\underline{\beta}\right) + \hat{\beta}'X'X\hat{\beta} + \underline{\beta}'\underline{V}^{-1}\underline{\beta}
\end{aligned}$$

Now, using the notation from part ii), we can write:

$$\beta'\left(X'X + \underline{V}^{-1}\right)\beta - 2\beta'\left(X'X\hat{\beta} + \underline{V}^{-1}\underline{\beta}\right) = \left(\beta-\bar{\beta}\right)' \bar{V}^{-1}\left(\beta-\bar{\beta}\right) - \bar{\beta}'\bar{V}^{-1}\bar{\beta}$$

So we have:

$$\left(\beta-\hat{\beta}\right)' X'X \left(\beta-\hat{\beta}\right) + \left(\beta-\underline{\beta}\right)' \underline{V}^{-1}\left(\beta-\underline{\beta}\right) = \left(\beta-\bar{\beta}\right)' \bar{V}^{-1}\left(\beta-\bar{\beta}\right) - \bar{\beta}'\bar{V}^{-1}\bar{\beta} + \hat{\beta}'X'X\hat{\beta} + \underline{\beta}'\underline{V}^{-1}\underline{\beta}$$

So we can rewrite Eq (1) as:

$$h^{\frac{n+\underline{\nu}-2}{2}} \exp\left(\frac{-\underline{\nu} h}{2\underline{h}}\right) \frac{\exp\left(\frac{-h}{2}\left[\hat{e}'\hat{e} - \bar{\beta}'\bar{V}^{-1}\bar{\beta} + \hat{\beta}'X'X\hat{\beta} + \underline{\beta}'\underline{V}^{-1}\underline{\beta}\right]\right) \det\left(h^{-1}\bar{V}\right)}{\det\left(h^{-1}\underline{V}\right)} \frac{\exp\left(-\frac{h}{2}\left[\left(\beta-\bar{\beta}\right)' \bar{V}^{-1}\left(\beta-\bar{\beta}\right)\right]\right)}{(2\pi)^{\frac{k}{2}} \det\left(h^{-1}\bar{V}\right)}$$

Now we can easily integrate β out since the the last term multiplying is the density of a normal distribution with mean $\bar{\beta}$ and covariance matrix $h^{-1}\bar{V}$

Therefore, we have:

$$\begin{aligned}
p(h \mid y) &= h^{\frac{n+\underline{\nu}-2}{2}} \exp\left(\frac{-\underline{\nu} h}{2\underline{h}}\right) \frac{\exp\left(\frac{-h}{2}\left[\hat{e}'\hat{e} - \bar{\beta}'\bar{V}^{-1}\bar{\beta} + \hat{\beta}'X'X\hat{\beta} + \underline{\beta}'\underline{V}^{-1}\underline{\beta}\right]\right) \det\left(h^{-1}\bar{V}\right)}{\det\left(h^{-1}\underline{V}\right)} \\
&\propto h^{\frac{n+\underline{\nu}-2}{2}} \exp\left(\frac{-h}{2}\left[\frac{\underline{\nu}}{\underline{h}} + \hat{e}'\hat{e} - \bar{\beta}'\bar{V}^{-1}\bar{\beta} + \hat{\beta}'X'X\hat{\beta} + \underline{\beta}'\underline{V}^{-1}\underline{\beta}\right]\right)
\end{aligned}$$

Comparing this to the kernel of a gamma distribution $\Gamma\left(\bar{\sigma}^2, \bar{\nu}\right) = \kappa h^{\frac{\bar{\nu}-2}{2}} \exp\left(\frac{-h\bar{\nu}}{2\bar{\sigma}^2}\right)$ we find:

$$\begin{aligned}
\bar{\nu} &= n + \underline{\nu} \\
\bar{\sigma}^2 &= \frac{\bar{\nu}}{\frac{\underline{\nu}}{\underline{h}} + \hat{e}'\hat{e} - \bar{\beta}'\bar{V}^{-1}\bar{\beta} + \hat{\beta}'X'X\hat{\beta} + \underline{\beta}'\underline{V}^{-1}\underline{\beta}}
\end{aligned}$$

Problem 2

Part a

The likelihood for the model is given by:

$$L\left(y, \theta_1, \theta_2, \beta_1, \beta_2, \sigma_1^2, \sigma_2^2, \lambda, x_t\right) = \Pr\left[y \mid \theta_1, \theta_2, \beta_1, \beta_2, \sigma_1^2, \sigma_2^2, \lambda, x_t\right] =$$

$$\prod_{t=1}^{\lambda} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{-\frac{(y_t-\theta_1-\theta_2 x_t)^2}{2\sigma_1^2}\right\} \prod_{t=\lambda+1}^{T} \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left\{-\frac{(y_t-\beta_1-\beta_2 x_t)^2}{2\sigma_2^2}\right\} =$$

$$\left(2\pi\sigma_1^2\right)^{\frac{-\lambda}{2}} \exp\left(\frac{-1}{2\sigma_1^2}\sum_{t=1}^{\lambda}(y_t-\theta_1-\theta_2 x_t)^2\right)\left(2\pi\sigma_2^2\right)^{-\frac{T-\lambda}{2}} \exp\left(\frac{-1}{2\sigma_2^2}\sum_{t=\lambda+1}^{T}(y_t-\beta_1-\beta_2 x_t)^2\right)$$

Part b)

In order to estimate the parameters using the Gibbs sampler, we need to find some conditional distributions. Lets start with θ. Using Bayes' Rule:

$$\Pr\left(\theta \mid \beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda,y_t,x_t\right) = \frac{\Pr\left[y \mid \theta_1,\theta_2,\beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda,x_t\right]\Pr\left[\theta_1,\theta_2,\beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda \mid x_t\right]}{\Pr\left[y_t,\beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda \mid x_t\right]}$$

Since we are interested in obtaining the kernel of the distribution of $\theta \mid \beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda,y_t,x_t$, we can neglect the term $\Pr\left[y_t,\beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda \mid x_t\right]$ and write:

$$\Pr\left(\theta \mid \beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda,y_t,x_t\right) \propto \Pr\left[y \mid \theta_1,\theta_2,\beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda,x_t\right]\Pr\left[\theta_1,\theta_2,\beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda \mid x_t\right]$$

The only two terms that matter for us in the expression above is $\Pr\left[y \mid \theta_1,\theta_2,\beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda,x_t\right]$ and the prior for θ. We can write:

$$\Pr\left(\theta \mid \beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda,y_t,x_t\right) \propto \exp\left(\frac{-1}{2\sigma_1^2}\sum_{t=1}^{\lambda}(y_t-\theta_1-\theta_2 x_t)^2\right)\exp\left(\frac{-1}{2}\left(\theta-\mu_\theta\right)' V_\theta^{-1}\left(\theta-\mu_\theta\right)\right)$$

Now let:

$$X_\theta = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_\lambda \end{bmatrix}$$

$$Y_\theta = \begin{bmatrix} y_1 \\ \vdots \\ y_\lambda \end{bmatrix}$$

$$\hat{\theta} = \left(X_\theta' X_\theta\right)^{-1} X_\theta' Y_\theta$$

Using the same idea we used in problem 1, we can rewrite $\exp\left(\frac{-1}{2\sigma_1^2}\sum_{t=1}^{\lambda}(y_t-\theta_1-\theta_2 x_t)^2\right)\exp\left(\frac{-1}{2}\left(\theta-\mu_\theta\right)' V_\theta^{-1}\left(\theta-\mu_\theta\right)\right)$ and find that:

$$\theta \mid \beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda,y_t,x_t \sim \mathcal{N}\left(\bar{\theta},\overline{V_\theta}\right)$$

where:

$$\overline{V_\theta} = \left(\frac{X_\theta' X_\theta}{\sigma_1^2} + V_\theta^{-1}\right)^{-1}$$

$$\bar{\theta} = \left(\frac{X_\theta' X_\theta}{\sigma_1^2} + V_\theta^{-1}\right)^{-1}\left(\frac{X_\theta' X_\theta}{\sigma_1^2}\hat{\theta} + V_\theta^{-1}\mu_\theta\right)$$

Now we will find $\Pr\left(\beta \mid \theta,\sigma_1^2,\sigma_2^2,\lambda,y_t,x_t\right)$. This is symmetric to the case $\Pr\left(\theta \mid \beta_1,\beta_2,\sigma_1^2,\sigma_2^2,\lambda,y_t,x_t\right)$. So we can define:

$$X_\beta = \begin{bmatrix} 1 & x_{\lambda+1} \\ \vdots & \vdots \\ 1 & x_T \end{bmatrix}$$
$$Y_\beta = \begin{bmatrix} y_{\lambda+1} \\ \vdots \\ y_T \end{bmatrix}$$
$$\hat{\beta} = \left(X'_\beta X_\beta\right)^{-1} X'_\beta Y_\beta$$

and we have:

$$\beta \mid \theta, \sigma_1^2, \sigma_2^2, \lambda, y_t, x_t \sim \mathcal{N}\left(\bar{\beta}, \overline{V_\beta}\right)$$

where:

$$\overline{V_\beta} = \left(\frac{X'_\beta X_\beta}{\sigma_2^2} + V_\beta^{-1}\right)^{-1}$$
$$\bar{\beta} = \left(\frac{X'_\beta X_\beta}{\sigma_2^2} + V_\beta^{-1}\right)^{-1}\left(\frac{X'_\beta X_\beta}{\sigma_2^2}\hat{\beta} + V_\beta^{-1}\mu_\beta\right)$$

We now find: $\sigma_1^2 \mid \theta, \beta, \sigma_2^2, \lambda, y_t, x_t$. I will call $\sigma_1^2 = h$ to avoid confusion:
First,

$$IG\left(a_1, a_2\right) \propto h^{-(a_1+1)} \exp\left(-\frac{1}{ha_2}\right)$$

$$\Pr\left(h \mid \theta, \beta, \sigma_2^2, \lambda, y_t, x_t\right) \propto h^{-\lambda/2} \exp\left(\frac{-1}{2h}\sum_{t=1}^{\lambda}\left(y_t - \theta_1 - \theta_2 x_t\right)^2\right) h^{-(a_1+1)} \exp\left(-\frac{1}{ha_2}\right)$$
$$= h^{-\left(a_1 + \frac{\lambda}{2} + 1\right)} \exp\left(\frac{-1}{h}\left[\frac{1}{a_2} + \frac{1}{2}\sum_{t=1}^{\lambda}\left(y_t - \theta_1 - \theta_2 x_t\right)^2\right]\right)$$

so $\sigma_1^2 \mid \theta, \beta, \sigma_2^2, \lambda, y_t, x_t$ also follows an inverse gamma distribution $IG\left(\overline{a_1}, \overline{a_2}\right)$, with:

$$\overline{a_1} = a_1 + \frac{\lambda}{2}$$
$$\overline{a_2} = \frac{1}{\frac{1}{a_2} + \frac{1}{2}\sum_{t=1}^{\lambda}\left(y_t - \theta_1 - \theta_2 x_t\right)^2}$$

Given the symmetry,

$$\sigma_2^2 \mid \theta, \beta, \sigma_2^2, \lambda, y_t, x_t \sim IG\left(\overline{b_1}, \overline{b_2}\right)$$

where

$$\overline{b_1} = b_1 + \frac{T - \lambda}{2}$$
$$\overline{b_2} = \frac{1}{\frac{1}{b_2} + \frac{1}{2}\sum_{t=\lambda+1}^{T}\left(y_t - \beta_1 - \beta_2 x_t\right)^2}$$

Now the last posterior we need to find is: $\lambda \mid \theta, \beta, \sigma_1^2, \sigma_2^2, y_t, x_t$. I will calculate the probability of λ assuming a specific value i given $\theta, \beta, \sigma_1^2, \sigma_2^2, y_t, x_t$. Again, using Baye's Rule:

$$\Pr\left[\lambda = i \mid \theta, \beta, \sigma_1^2, \sigma_2^2, y_t, x_t\right] \propto \Pr\left[y \mid \theta_1, \theta_2, \beta_1, \beta_2, \sigma_1^2, \sigma_2^2, \lambda = i, x_t\right] \Pr\left[\lambda = i\right]$$
$$= \left(2\pi\sigma_1^2\right)^{\frac{-i}{2}} \exp\left(\frac{-1}{2\sigma_1^2}\sum_{t=1}^{i}\left(y_t - \theta_1 - \theta_2 x_t\right)^2\right)\left(2\pi\sigma_2^2\right)^{-\frac{T-i}{2}} \exp\left(\frac{-1}{2\sigma_2^2}\sum_{t=i+1}^{T}\left(y_t - \beta_1 - \beta_2 x_t\right)^2\right)\frac{1}{T-1}$$

Since $\sum_{i=1}^{T-1} \Pr\left[\lambda = i \mid \theta, \beta, \sigma_1^2, \sigma_2^2, y_t, x_t\right] = 1$ we can find the exact value of $\Pr\left[\lambda = i \mid \theta, \beta, \sigma_1^2, \sigma_2^2, y_t, x_t\right]$ as:

$$\Pr\left[\lambda = i \mid \theta, \beta, \sigma_1^2, \sigma_2^2, y_t, x_t\right] = \frac{\left(2\pi\sigma_1^2\right)^{\frac{-i}{2}} \exp\left(\frac{-1}{2\sigma_1^2} \sum_{t=1}^{i} \left(y_t - \theta_1 - \theta_2 x_t\right)^2\right) \left(2\pi\sigma_2^2\right)^{-\frac{T-i}{2}} \exp\left(\frac{-1}{2\sigma_2^2} \sum_{t=i+1}^{T} \left(y_t - \beta_1 - \beta_2 x_t\right)^2\right) \frac{1}{T-1}}{\sum_{i=1}^{T-1} \left(2\pi\sigma_1^2\right)^{\frac{-i}{2}} \exp\left(\frac{-1}{2\sigma_1^2} \sum_{t=1}^{i} \left(y_t - \theta_1 - \theta_2 x_t\right)^2\right) \left(2\pi\sigma_2^2\right)^{-\frac{T-i}{2}} \exp\left(\frac{-1}{2\sigma_2^2} \sum_{t=i+1}^{T} \left(y_t - \beta_1 - \beta_2 x_t\right)^2\right) \frac{1}{T-1}} \tag{2}$$

So our Gibbs sampler procedure for estimating the parameters can be described as:

1. Start with a initial guess for $\lambda, \sigma_1^2, \sigma_2^2$. I chose the mean values of the priors for these parameters.
2. Draw θ from a $\mathcal{N}\left(\bar{\theta}, \overline{V_\theta}\right)$
3. Draw β from $\mathcal{N}\left(\bar{\beta}, \overline{V_\beta}\right)$
4. Given our draw for θ, draw σ_1^2 from $IG\left(\overline{a_1}, \overline{a_2}\right)$
5. Given our draw for β, draw σ_2^2 from $IG\left(\overline{b_1}, \overline{b_2}\right)$
6. Given the draws we got from step 2 to 6, draw λ from a discrete distribution, assuming values $i \in \{1, 2 \ldots T-1\}$ with probabilities given by Eq. 2.
7. Repeat steps 2-6 S times.
8. Discard the first B iterations

The estimate for each parameter is just the average of the draws for each parameter. The following table shows the estimates:

Table 1: Parameters Estimates

Parameter	Estimate	True Value
θ_1	1.97	2
θ_2	0.98	1
β_1	1.43	1.5
β_2	0.81	0.8
σ_1^2	0.20	0.2
σ_2^2	0.47	0.5
λ	85.70	85

Problem 3

First, there is an identification problem in the Probit model. Our model is $y_i^* = x_i'\beta + \sigma e_i$, with $e_i \sim \mathcal{N}(0, 1)$. We observe $y_i = 1$ if $y_i^* > 0$. So the probability of seeing $y_i = 1$ is given by:

$$\begin{aligned}\Pr\left[x_i'\beta + \sigma e_i > 0\right] &= \Pr\left[e_i > \frac{-x_i'\beta}{\sigma}\right] \\ &= \Phi\left[\frac{x_i'\beta}{\sigma}\right]\end{aligned}$$

where

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} \mathrm{d}t$$

It is clear that if we multiply both β and σ by the same factor, say 2, we will have the model. In order to fix the identification problem, we will set $\sigma = 1$.
Now we will derive some of the posteriors we will use in the Gibbs sampling.
Note that:

$$\begin{aligned}\Pr\left[y_i^* \leq t \mid \beta, y_i = 1\right] &= \Pr\left[y_i^* \leq t \mid \beta, y_i^* > 0\right] \\ &= \frac{\Pr\left[y_i^* \leq t, y_i^* > 0 \mid \beta\right]}{\Pr\left[y_i^* > 0 \mid \beta\right]} \\ &= \frac{\Phi\left(t - x_i'\beta\right) - \Phi\left(x_i'\beta\right)}{\Phi\left(x_i'\beta\right)}\end{aligned}$$

and similarly:

$$\Pr\left[y_i^* \leq t \mid \beta, y_i = 0\right] = \Pr\left[y_i^* \leq t \mid \beta, y_i^* \leq 0\right] = \frac{\Pr\left[y_i^* \leq t, y_i^* \leq 0 \mid \beta\right]}{\Pr\left[y_i^* \leq 0 \mid \beta\right]} = \frac{\Phi\left(t - x_i'\beta\right)}{1 - \Phi\left(x_i'\beta\right)}$$

We can conclude that

$$y_i^* \mid \beta, y_i \sim \begin{cases} TN_{(0,\infty)}\left(x_i'\beta, 1\right) & y_i = 1 \\ TN_{(-\infty,0)}\left(x_i'\beta, 1\right) & y_i = 0 \end{cases}$$

The other posterior we will need is $\beta \mid y^*, y$. Note first that, once we know y^* we know the value of y, so we can focus on $\beta \mid y^*$. Assuming a non informative prior for β, we have:

$$\Pr\left[\beta \mid y^*\right] \propto \Pr\left[y^* \mid \beta\right]$$

and

$$\begin{aligned}\Pr\left[y^* \mid \beta\right] &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\left(y_i^* - x_i'\beta\right)^2}{2}\right) \\ &= (2\pi)^{-\frac{n}{2}} \exp\left(\frac{-\sum_{i=1}^{n}\left(y_i^* - x_i'\beta\right)^2}{2}\right)\end{aligned}$$

Let $\hat{\beta}$=$\left(X'X\right)^{-1} X'Y^*$, then we have:

$$\begin{aligned}\sum_{i=1}^{n}\left(y_i^* - x_i'\beta\right)^2 &= \left(Y^* - X\beta\right)'\left(Y^* - X\beta\right) \\ &= \left(Y^* - X\hat{\beta} - X\left(\beta - \hat{\beta}\right)\right)'\left(Y^* - X\hat{\beta} - X\left(\beta - \hat{\beta}\right)\right) \\ &= \left(Y^* - X\hat{\beta}\right)'\left(Y^* - X\hat{\beta}\right) + \left(\beta - \hat{\beta}\right)' X'X\left(\beta - \hat{\beta}\right)\end{aligned}$$

So we can rewrite $\Pr\left[y^* \mid \beta\right]$ as:

$$\Pr\left[y^* \mid \beta\right] \propto \exp\left(\frac{-\left(\beta - \hat{\beta}\right)' X'X\left(\beta - \hat{\beta}\right)}{2}\right)$$

which is the kernel of a normal density with mean $\hat{\beta}$ and covariance matrix $\left(X'X\right)^{-1}$.
So we process as follows in the Gibbs sampling:

1. Given an initial value for β we generate y_i^* by sampling from the truncated normal, as we saw above. If $y_i = 1$ then we sample from $TN_{(0,\infty)}\left(x_i'\beta, 1\right)$, if $y_i = 0$ then we sample from $TN_{(-\infty,0)}\left(x_i'\beta, 1\right)$.

2. Given y_i^*, we sample β from a Normal with mean $\hat{\beta} = \left(X'X\right)^{-1} X'Y^*$ and variance $\left(X'X\right)^{-1}$.

3. With this new β, we generate $y_i^*$by sampling from the truncated normal, as we saw above. If $y_i = 1$ then we sample from $TN_{(0,\infty)}(x_i'\beta, 1)$, if $y_i = 0$ then we sample from $TN_{(-\infty,0)}(x_i'\beta, 1)$.

We repeat steps 2 and 3 above $S + B$ times where B is the burn in.
In order to estimate $\Pr[y_i = 1 \mid x_i = 2]$ we use the following:

$$\begin{aligned}\Pr[y_i = 1 \mid x_i = 2] &= \Pr[y_i^* > 0 \mid x_i = 2] \\ &= \Pr[\beta_0 + 2\beta_1 + e_i > 0] \\ &= \Pr[e_i > -\beta_0 - 2\beta_1] \\ &= \Phi(\beta_0 + 2\beta_1)\end{aligned}$$

So in each iteration of the Gibbs sampler we compute $\Phi(\beta_0 + 2\beta_1)$ given the the samples for β we have in that iteration.
The following table reproduces the results.

Table 2: Parameter Estimates

Variable	Posterior Mean	Posterior St. Dev.
β_0	-0.0087	0.1303
β_1	0.4315	0.1322
$\Phi(\beta_0 + 2\beta_1)$	0.7936	0.0815