

## Econ 512

Fall 2019

### Homework 5 – Binary Choice MLE with Random Coefficients

Due 12/02/2019

Consider the following binary discrete choice model for panel data:

$$Y_{it} = I(\beta_i X_{it} + \gamma Z_{it} + u_i + \epsilon_{it} > 0),$$

where  $X_{it}$  and  $Z_{it}$  are scalar regressors for a group of  $i = 1, \dots, N$  individuals and  $t = 1, \dots, T$  time periods.  $N = 100$  and  $T = 20$ . The coefficients  $\beta_i$  and  $u_i$  are person specific and are modeled as draws from a bivariate normal distribution:

$$\begin{bmatrix} \beta_i \\ u_i \end{bmatrix} \sim N(\mu, \Sigma), \quad \text{where} \quad \mu = \begin{bmatrix} \beta_0 \\ u_0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_\beta & \sigma_{\beta u} \\ \sigma_{\beta u} & \sigma_u \end{bmatrix}.$$

Assume  $\epsilon_{it}$  follows standard logistic distribution, i.e.  $F(\epsilon) = (1 + e^{-\epsilon})^{-1}$ , and then a single contribution to the likelihood function from individual  $i$  is

$$L_i(\gamma \mid \beta_i, u_i) = \prod_{t=1}^T F(\beta_i X_{it} + \gamma Z_{it} + u_i)^{Y_{it}} [1 - F(\beta_i X_{it} + \gamma Z_{it} + u_i)]^{1-Y_{it}}.$$

To construct the likelihood function for the data set of NT observations we have to integrate over the joint distribution of  $(\beta_i, u_i)$ . The likelihood function for the data set is:

$$L(\gamma, \mu, \Sigma) = \prod_{i=1}^N \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L_i(\gamma \mid \beta_i, u_i) \phi(\beta_i, u_i \mid \mu, \Sigma) d\beta_i du_i,$$

where  $\phi(\cdot \mid \mu, \Sigma)$  is the joint density function of bivariate normal distribution  $N(\mu, \Sigma)$ . Of course, it will be numerically more convenient to work with the log-likelihood function. The data set `hw5.mat` contains 20 x 100 matrices of the variables  $X$ ,  $Z$ , and  $Y$ .

1. Assume  $u_i = 0 \forall i$  (ie. take  $u_i$  out of the model, so that  $u_0 = \sigma_u = \sigma_{u\beta} = 0$ ). Use Gaussian Quadrature using 20 nodes to calculate the log-likelihood function when  $\beta_0 = 0.1$ ,  $\sigma_\beta = 1$ , and  $\gamma = 0$ .
2. Now use Monte Carlo Methods using 100 nodes to calculate the log-likelihood function.
3. Maximize (or minimize the negative) log-likelihood function with respect to the parameters using both integration techniques above. Use Matlab's `fmincon` without a supplied derivative to max (min) your objective function.
4. Now allow  $u_0 \neq 0$ , so allow the parameters  $\sigma_u$  and  $\sigma_{\beta u}$  to be non-zero. Maximize the log-likelihood function, estimating all of the parameters, using Monte Carlo methods.
5. For each estimation, report the starting value, argmax, and maximized value of the log-likelihood function.

(Hint: the matlab function “chol” may come in handy for simulating from the joint density.)