## Question 3: (Fully) Parametric, Structural Estimation

Under the assumption that $\epsilon$ and $\xi$ follow a bivariate normal distribution

$$\binom{\epsilon}{\xi} \sim \mathcal{N}\left(\left[\begin{array}{l}0 \\ 0\end{array}\right],\left[\begin{array}{cc}\sigma_\epsilon^2 & \rho \sigma_\epsilon \sigma_\xi \\ \rho \sigma_\epsilon \sigma_\xi & \sigma_\xi^2\end{array}\right]\right)$$

**(a) Estimate the model using full information maximum likelihood (that is, deriving a likelihood function in which we are estimating the utility function and the wage equation jointly).**

In this exercise we will use the estimation procedure in the slides of Static Labor Participation Estimation. In that specific slide we have:

"… the likelihood can be written in the following way:

$$\mathcal{L}(\theta ; \text { data })=\prod_{i=1}^N\left(\operatorname{Pr}\left(P_i=1, w=w_i \mid y_i, \mathbf{z}, \mathbf{X}\right)\right)^{P_i}\left(\operatorname{Pr}\left(P_i=0 \mid y_i, \mathbf{z}, \mathbf{X}\right)\right)^{1-P_i}$$

where:

$$\operatorname{Pr}\left(P_i=0 \mid y_i, \mathbf{z}, \mathbf{X}\right)=\Phi\left(\frac{\mathbf{X}_i^{\prime} \alpha+\beta y_i-\mathbf{z}^{\prime} \gamma}{\sigma_\nu}\right)$$

with $\sigma_\nu^2=\sigma_\epsilon^2+\sigma_\xi^2-2 \rho \sigma_\epsilon \sigma_\xi$, and

$$\begin{aligned}
& \operatorname{Pr}\left(P_i=1, w=w_i \mid y_i, \mathbf{z}, \mathbf{X}\right)=\frac{1}{\sigma_\xi} \phi\left(\frac{w_i-\mathbf{z}^{\prime} \gamma}{\sigma_\nu}\right) \\
& \times \Phi\left(\frac{w_i-\mathbf{X}_i^{\prime} \alpha-\beta y_i-\rho \frac{\sigma_\epsilon}{\sigma_\xi}\left[w_i-\mathbf{z}_i^{\prime} \gamma\right]}{\sigma_\epsilon \sqrt{1-\rho^2}}\right)
\end{aligned}$$

" In our context, if we let $X_i^{\prime}=[1, age_i, nchild_i]^{\prime}$, $y_i$ is non-labor income and $z_i^{\prime}=[age_i, education_i]$,and $\Phi(.), \phi(.)$ represents the cdf and density of a standard normal.

We do it by first defining the log_likelihood for each sample as we see below:

```
# 3)Under the assumption that epsilon and  follow a bivariate normal distribution
# 3-a) full information maximum likelihood --------------------------------
#  Estimate the model using full information maximum likelihood
# (that is, deriving a likelihood function in which we are estimating
# the utility function and the wage equation jointly).
```

```
# Z = educ and age
# X = const, age , nchild
# y = non labor income

# Loglikelihood P-N First sample of 1000bs
log_likelihood_n1 <- function(par){

  # Dictionary
  # par[1] = alpha_1
  # par[2] = alpha_2
  # par[3] = alpha_3
  # par[4] = beta
  # par[5] = gamma_1
  # par[6] = gamma_2
  # par[7] = sigma_e
  # par[8] = sigma_xi
  # par[9] = rho


  # sigma_v
  sigma_v <- sqrt(par[7]^2 + par[8]^2 - 2*par[9]*par[7]*par[8])

  # Here we define the Argument of Pr(P=0)
  arg_pr_0 <- (par[1] + par[2]*X_n1[,2] + par[3]*X_n1[,3] + par[4]*Y_n1 - par[5]*Z_n1[,1]

  # Define we properly define Pr(0)
  pr_0 <- pnorm(arg_pr_0)
  pr_0 <- exp(pr_0) / (1 + exp(pr_0))

  # Here we define Arg 1 of Pr(1), the argument of the density normal
  arg_pr_1 <- (W_n1 - par[5]*Z_n1[,1] - par[6]*Z_n1[,2]) / sigma_v

  # Here we define the Arg 2 of Pr(1), the argument of the cdf normal
  arg_pr_2 <-  W_n1 - par[1] - par[2]*X_n1[,2]- par[3]*X_n1[,3] -par[4]*Y_n1
  arg_pr_2 <-  arg_pr_2 -(par[9]*par[7]/sigma_v)*(W_n1-par[5]*Z_n1[,1]-par[6]*Z_n1[,2])
  arg_pr_2 <- arg_pr_2 / (par[7]*sqrt(1-par[9]^2))

  # Values of Pr(1)
  pr_1 <- (1/par[8]) * dnorm(arg_pr_1) * rnorm(arg_pr_2)
  pr_1 <- exp(pr_1)/(1 + exp(pr_1))
```

```
  # Objective function
  log_like = sum(L_n1*log(pr_1) + (1 - L_n1)*(log(pr_0)))
  return(log_like)
}
```

Now we define the same function but for the sample of 5000 obs.

```
log_likelihood_n2 <- function(par){

  # Dictionary
  # par[1] = alpha_1
  # par[2] = alpha_2
  # par[3] = alpha_3
  # par[4] = beta
  # par[5] = gamma_1
  # par[6] = gamma_2
  # par[7] = sigma_e
  # par[8] = sigma_xi
  # par[9] = rho


  # sigma_v
  sigma_v <- sqrt(par[7]^2 + par[8]^2 - 2*par[9]*par[7]*par[8])

  # Arg pr_0
  arg_pr_0 <- (par[1] + par[2]*X_n2[,2] + par[3]*X_n2[,3] + par[4]*Y_n2 - par[5]*Z_n2[,1]

  # Define Pr(0)
  pr_0 <- pnorm(arg_pr_0)
  pr_0 <- exp(pr_0) / (1 + exp(pr_0))

  # Arg 1 of pR(1)
  arg_pr_1 <- (W_n2 - par[5]*Z_n2[,1] - par[6]*Z_n2[,2]) / sigma_v

  # Arg 2 de pr2
  arg_pr_2 <-  W_n2 - par[1] - par[2]*X_n2[,2]- par[3]*X_n2[,3] -par[4]*Y_n2
  arg_pr_2 <-  arg_pr_2 -(par[9]*par[7]/sigma_v)*(W_n2-par[5]*Z_n2[,1]-par[6]*Z_n2[,2])
  arg_pr_2 <- arg_pr_2 / (par[7]*sqrt(1-par[9]^2))

  # Pr(1)
  pr_1 <- (1/par[8]) * dnorm(arg_pr_1) * rnorm(arg_pr_2)
```

```
  pr_1 <- exp(pr_1)/(1 + exp(pr_1))

  # Objective
  log_like = sum(L_n2*log(pr_1) + (1 - L_n2)*(log(pr_0)))
  return(log_like)
}
```

Finally with following commands we optimize the log likelihood function in order to find the parameters. For initial values we could use the *lucky guess* which are values we previously estimated and know are near the optimum or we can start by using a random draw of number between -100 and 100 and keep iterating the process until we converge to the desired point.

```
# Guesses
lucky_guess <- c(9.156, 0.858, -0.609, 1.265, 1232.229, 60.274,100, 100, 0)
random_guess <- sample(-100:100, 9)


#  See if has correct number of parameters =9
length(lucky_guess)
length(random_guess)

# Optim with luck_guess
optim(lucky_guess, log_likelihood_n1, method = "Nelder-Mead")
optim(lucky_guess, log_likelihood_n2, method = "Nelder-Mead")

# Optim with random_guess
optim(random_guess, log_likelihood_n1, method = "Nelder-Mead")
optim(random_guess, log_likelihood_n2, method = "Nelder-Mead")
```

Table 8: Estimation Parametric-Structural

| Parameter | 5000 Obs. | 1000 Obs |
|---|---|---|
| Con$\alpha_1$ | 56.35703 | 49.70609 |
| Age $\alpha_2$ | 47.66128 | -26.41636 |
| nchild $\alpha_3$ | 44.87264 | 34.93082 |
| Non-Labor Income $\beta$ | -219.79045 | -149.24950 |
| Age $\gamma_1$ | 1277.57623 | 1322.04343 |
| education $\gamma_2$ | 113.59084 | 145.01691 |
| $\sigma_\epsilon$ | 145.38495 | 122.68997 |
| $\sigma_\chi$ | 149.16555 | 103.24569 |
| $\rho$ | 40.25317 | 103.24569 |

| Parameter | 5000 Obs. | 1000 Obs |
|---|---|---|
| Optim Value | -3465.75 | -693.1539 |

**(b) Compare your results with those obtained in Question 2. Discuss.**