Spring 1998 John Rust
Economics 551b 37 Hillhouse, Rm. 27

PROBLEM SET 1
Nonlinear Estimation of Binary Choice Models

QUESTION 1 Extract data in file `data3.asc` in the

`pub/John_Rust/courses/econ551/regression/`

directory on `gemini.econ.yale.edu` (either ftp to `gemini.econ.yale.edu` and login as "anonymous" and `cd pub/John_Rust/courses/econ551/regression` and `get data1.asc` or click on the hyperlink in the html version of this document). This data file contains $n = 3000$ *IID* observations (y_i, x_i) that I generated from the binary probability model:

$$y = \begin{cases} 1 & \text{with probability} \quad \Psi(x,\theta) \\ 0 & \text{with probability} \quad 1-\Psi(x,\theta) \end{cases} \tag{1}$$

where $\Psi(x,\theta)$ is some parametric model of the conditional probability of the binary variable y given x, i.e. $\Psi(x,\theta) = P\{y = 1|x,\theta\}$. Two standard models for Ψ are the *logit* and *probit* models. In the logit model we have

$$\Psi(x,\theta) = \frac{\exp\{x'\theta\}}{1+\exp\{x'\theta\}} \tag{2}$$

and in the probit mode we have

$$\Psi(x,\theta) = \Phi(x'\theta), \tag{3}$$

where $\Phi(x)$ is the standard normal CDF, i.e.

$$\Phi(x) = \int_{-\infty}^{x} \phi(y)dy$$

where

$$\phi(y) = \frac{1}{\sqrt{2\pi}}\exp\{-y^2/2\}.$$

More generally, Ψ could take the form

$$\Psi(x,\theta) = F(x'\theta)$$

where F is an arbitrary continuous CDF.

1. Show that versions of the logit and probit models can be derived from an underlying *random utility model* where a decision maker has utility function of the form:

   $$U(y,x,\epsilon,\theta) = u(y,x,\theta) + \epsilon(y), \quad y = 0,1$$

   and takes action $y = 1$ if $u(1,x,\theta) + \epsilon(1) > u(0,x,\theta) + \epsilon(0)$ and takes action $y = 0$ if $u(1,x,\theta) + \epsilon(1) \le u(0,x,\theta) + \epsilon(0)$. Derive the implied *choice probability* $\Psi(x,\theta) = P\{y|x,\theta\}$ in the case where $\{\epsilon(0),\epsilon(1)\}$ is a bivariate normal random vector with $E\{\epsilon(i)\} = 0$, $\text{var}\{\epsilon(i)\} = 1/2$ and $\text{cov}(\epsilon(1),\epsilon(0)\} = 0$ and $u(1,x,\theta) = x'\theta$ and $u(0,x,\theta) = 0$. What is the form of $\Psi(x,\theta)$ in the general case when $\{\epsilon(0),\epsilon(1)\}$ has an unrestricted bivariate normal distribution with mean vector μ and covariance matrix Ω? If the utility function includes a constant term, i.e. $u(1,x,\theta) = \theta_0 + x'\theta_1$ are the θ, μ and Ω parameters all separately identified if we only have access to data on (y,x) pairs?

2. Derive the form of the choice probability under the same assumptions are part 1 above but when $\{\epsilon(0), \epsilon(1)\}$ has a bivariate *Type I extreme value distribution* by doing problem 7 of the 1997 Econ 551 problem set 3. By doing this you will have derived the binary logit model from first principles.

3. Using the artificially generated data in

   `pub/John_Rust/courses/econ551/regression/data3.asc`

   compute maximum likelihood estimates of the parameters $(\theta_0, \theta_1, \theta_2, \theta_3)$ of the logit and probit specifications given in equations (2) and (3) above, where $x'\theta$ is given by:

$$x'\theta = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3.$$

4. Is it possible to consistently estimate θ by doing nonlinear least squares estimation of the nonlinear regression formulation of the binary probability model

$$y = \Psi(x, \theta) + \eta \tag{4}$$

   instead of doing maximum likelihood? If so, provide a proof of the consistency of the NLLS estimator. If not, provide a counterexample showing that the NLLS estimator is inconsistent.

5. Estimate both the probit and logit specifications by nonlinear least squares as suggested in part (4). How do the parameter estimates and standard errors compare to the maximum likelihood estimates computed in part 3?

6. Is there any problem of heteroscedasticity in the nonlinear regression formulation of the problem in (4)? If so, derive the form of the heteroscedasticity and, using the estimated "first stage" parameters from part 5 above, compute second stage "feasible generalized least squares" (FGLS) estimates of θ.

7. Are the FGLS estimates of θ consistent and asymptotically normally distributed (assuming the model is correctly specified)? If so, derive the asymptotic distribution of the FGLS estimator, and if not provide a counter example showing that the FGLS estimator is inconsistent or not asymptotically normally distributed. If you conclude that the FGLS estimator is asymptotically normally distributed, is it as efficient as the maximum likelihood estimator of θ? Explain your reasoning for full credit.

8. Is it possible to determine whether the data in the file `data3.asc` are generated from a logit or probit model? In answering this question, consider whether you could estimate $\Psi(x, \theta)$ nonparametrically via non-parametric regression. Is there any way you could use the nonparametric regression estimate of Ψ to help discriminate between the logit and probit specifications?