

Problem Set #11

Danny Edgel
Econ 710: Economic Statistics and Econometrics II
Spring 2021

April 18, 2021

Discussed and/or compared answers with Sarah Bass, Emily Case, Katherine Kwok, Michael Nattinger, and Alex Von Hafften

Exercise 25.1

The slope coefficients represent the difference in log odds between the two dependent variable outcomes generated by the independent variables. Thus, the slope coefficients will have the same absolute values in each specification, but different signs.

Exercise 25.3

Recall that $P(x) = P[Y = 1|X = x]$, where $Y \in \{0, 1\}$. Then, when we have the model:

$$Y = P(X) + e$$

It must be the case that e is bounded by 0 and 1, where:

$$e = \begin{cases} 1 - P(x), & Y = 1 \\ P(x), & Y = 0 \end{cases}$$

Where $Y = 1$ with probability $P(x)$ and $Y = 0$ with probability $1 - P(x)$. Now, we can find the conditional variance of e :

$$\begin{aligned} \text{Var}(e|X) &= \mathbb{E}[e^2|X] - \mathbb{E}[e|X]^2 = \mathbb{E}[e^2|X] \\ e^2 &= Y^2 - 2YP(x) + P(x)^2 \\ \mathbb{E}[e^2|X] &= \mathbb{E}[Y^2|X] - 2\mathbb{E}[Y|X]P(x) + P(x)^2 \\ \mathbb{E}[Y|X] &= P(x) \\ \mathbb{E}[Y^2|X] &= (1)P(x) + (0)(1 - P(x)) = P(x) \\ \Rightarrow \text{Var}(e|X) &= P(x) - P(x)^2 = P(x)(1 - P(x)) \end{aligned}$$

Exercise 25.9

The logit log-likelihood function is:

$$\ell_n(\beta) = \sum_{i=1}^n \log(\Lambda(Z_i' \beta))$$

Where

$$Z_i = \begin{cases} X_i, & X_i Y = 1 \\ -X_i, & Y = 0 \end{cases}$$

Which yields the following first-order condition:

$$\begin{aligned} \frac{d\ell_n(\beta)}{d\beta} &= 0 \\ \sum_{i=1}^n \left[\frac{Z_i \Lambda(Z_i' \beta) (1 - \Lambda(Z_i' \beta))}{\Lambda(Z_i' \beta)} \right] &= 0 \\ \sum_{i=1}^n [Z_i (1 - \Lambda(Z_i' \beta))] &= 0 \end{aligned}$$

Exercise 25.12

A nonlinear least squares (NLS) estimator is derived by finding the parameter β that solves:

$$\min_{\beta} \frac{1}{n} \sum_{i=1}^n (Y_i - m(X_i | \beta))^2$$

Where $m(X_i | \beta)$ is the expectation of Y with respect to X . Let $\lambda(x) = \phi(x)/\Phi(x)$. Then, in a probit model,

$$\begin{aligned} \mathbb{E}[Y_i | X_i = x] &= \mathbb{E}[Y_i | X_i = x, Y_i^* \geq 0] \Pr(Y_i^* \geq 0 | X_i = x) \\ &= (x' \beta + \mathbb{E}[e_i | X_i = x, Y_i^* \geq 0]) \Pr(x' \beta + e_i \geq 0) \\ &= \left(x' \beta + \sigma \lambda \left(\frac{x' \beta}{\sigma} \right) \right) \Pr(e_i \geq -x' \beta) \\ &= \left(x' \beta + \sigma \lambda \left(\frac{x' \beta}{\sigma} \right) \right) \Phi(x' \beta) \end{aligned}$$

Then, the probit estimator is the solution to:

$$\min_{\beta, \sigma} \frac{1}{n} \sum_{i=1}^n \left(Y_i - \left(x' \beta + \sigma \lambda \left(\frac{x' \beta}{\sigma} \right) \right) \Phi(x' \beta) \right)^2$$

Exercise 25.14

1. In this case, the response probability is given as the probability that Y is not censored (i.e., equal to 0):

$$\begin{aligned} Pr(Y > 0) &= Pr(Y^* > 0) = Pr(m(X) + e > 0) = Pr(e > -m(X)) \\ &= 1 - \Phi(-m(X)) \end{aligned}$$

Where $\Phi(\cdot)$ is the mean zero normal CDF with a standard deviation equal to $\sigma^2(X)$.

2. No. $\sigma^2(X)$ is not known, so we can only identify $m(X)/\sigma^2(X)$.
3. Following from (b), we can identify a scaled version of $m(X)$ by estimating a model that assumes that errors are normally distributed with variance 1.
4. No, because this model, which assumes heteroskedastic errors, uses the same identification as a model that assumes homoskedastic errors. In any case, we interpret the coefficients as of the model as if they have been scaled by the model's variance.

Exercise 25.15

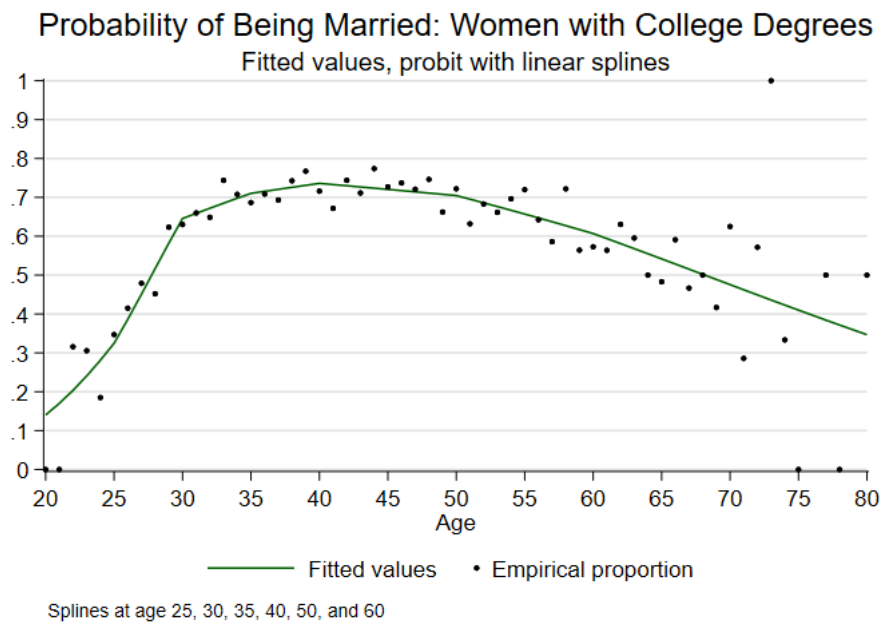
The results of the probit estimation are displayed below, with coefficients reported in the first column and average marginal effects reported in the second column for interpretation. The AME column provides the average change in probability of union membership caused by a change in the independent variable. This suggests that each year of age adds .042% to the odds that an individual belongs to a union, while the likelihood that Hispanic individuals to belong to a union is 1.58% lower, on average, than that of non-Black and non-Hispanic individuals (i.e. the omitted group). Finally, more educated men are less likely to belong in a union, likely reflecting lower unionization rates for white-collar jobs.

VARIABLES	(1) Probit	(2) AME
age	0.00793*** (0.00143)	0.000421*** (7.70e-05)
education	-0.0255*** (0.00631)	-0.00135*** (0.000338)
$\mathbb{1}\{Black\}$	-0.0541 (0.0600)	-0.00287 (0.00318)
$\mathbb{1}\{Hispanic\}$	-0.298*** (0.0565)	-0.0158*** (0.00304)
Constant	-1.954*** (0.109)	
Observations	29,140	29,140

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Exercise 25.17

The graph below displays the results of the estimation.¹ I chose to estimate the probability of marriage with a probit model that includes linear splines at the ages of 25, 30, 35, 50, 50, and 60. These dates were chosen based on intuition and anecdotal experience. However, compared to the empirical proportions shown in the chart below, they match the data well. There is a large increase in the share of women with college degrees who are married from age 20 to age 30, but the relationship levels off from 30 to 40, then reverses from 40 onward.



This result differs so substantially from the one displayed in figure 25.1 that I tested my exact code, but flipping the subsetting line from female to male just to make sure that there was not a critical error in my estimation. Sure enough, the result for men was similar to figure 25.1. The main difference is the reversal of the age-marriage relationship that begins around age 40, which does not exist for men but exists for women.

¹See the attached do file for the code that generated the graph.

Exercise 26.1

The model in question is:

$$P_j(x) = \frac{\exp(x'\beta_j)}{\sum_{\ell=1}^J \exp(x'\beta_\ell)}$$

The statement $0 \leq P_j(x) \leq 1$ is trivially clear by the fact that $P_j(x)$ is a fraction where the denominator is a sum of a weakly positive series that includes the numerator. Similarly, we can show that $\sum_{j=1}^J P_j(x) = 1$ as follows:

$$\sum_{j=1}^J P_j(x) = \sum_{j=1}^J \frac{\exp(x'\beta_j)}{\sum_{\ell=1}^J \exp(x'\beta_\ell)} = \frac{\sum_{j=1}^J \exp(x'\beta_j)}{\sum_{\ell=1}^J \exp(x'\beta_\ell)} = 1$$

Exercise 26.3

$$\begin{aligned} \frac{\partial}{\partial x} P_j(x) &= \frac{\partial}{\partial x} \frac{\exp(x'\beta_j)}{\sum_{\ell=1}^J \exp(x'\beta_\ell)} \\ &= \beta_j \left(\frac{\exp(x'\beta_j)}{\sum_{\ell=1}^J \exp(x'\beta_\ell)} \right) + \sum_{\ell=1}^J \left(-\beta_\ell \exp(x'\beta_\ell) \frac{\exp(x'\beta_j)}{\left(\sum_{\ell=1}^J \exp(x'\beta_\ell) \right)^2} \right) \\ &= \beta_j P_j(x) - \sum_{\ell=1}^J \beta_\ell \left(\frac{\exp(x'\beta_\ell)}{\sum_{\ell=1}^J \exp(x'\beta_\ell)} \right) \left(\frac{\exp(x'\beta_j)}{\sum_{\ell=1}^J \exp(x'\beta_\ell)} \right) \\ &= \beta_j P_j(x) - \sum_{\ell=1}^J \beta_\ell P_\ell(x) P_j(x) = P_j(x) \left(\beta_j - \sum_{\ell=1}^J \beta_\ell P_\ell(x) \right) \end{aligned}$$

Exercise 26.7

The marginal effect of the conditional logit model is:

$$\delta_{j\ell}(w, x) = \frac{\partial}{\partial x_\ell} P_j(w, x) = \begin{cases} \gamma P_j(w, x) (1 - P_j(w, x)), & \ell = j \\ -\gamma P_j(w, x) P_\ell(w, x), & \ell \neq j \end{cases}$$

Then, the average marginal effect is:

$$AME_{j\ell} = \mathbb{E}[\delta_{j\ell}(W, X)]$$

Which can be estimated as:

$$\widehat{AME}_{j\ell} = \begin{cases} \hat{\gamma} \sum_{i=1}^n \hat{P}_j(w, x) (1 - \hat{P}_j(w, x)), & \ell = j \\ -\hat{\gamma} \sum_{i=1}^n \hat{P}_j(w, x) \hat{P}_\ell(w, x), & \ell \neq j \end{cases}$$

Exercise 26.8

From equation (26.1), we have:

$$P_j(x) = \frac{\exp(x'\beta_j)}{\sum_{\ell=1}^J \exp(x'\beta_\ell)}$$

Then, we can solve:

$$\frac{P_j(W, X|\theta)}{P_\ell(W, X|\theta)} = \frac{\frac{\exp(W'\beta_j + X'_j\gamma)}{\sum_{k=1}^J \exp(W'\beta_k + X'_k\gamma)}}{\frac{\exp(W'\beta_\ell + X'_\ell\gamma)}{\sum_{k=1}^J \exp(W'\beta_k + X'_k\gamma)}} = \frac{\exp(W'\beta_j + X'_j\gamma)}{\exp(W'\beta_\ell + X'_\ell\gamma)}$$