Econ 140 - Spring 2016
Section 10

GSI: Fenella Carpena

April 13, 2016

1 Linear Probability Model (LPM)

- **Specification.**

- **Estimation Method.**

- **Example.** Consider the results of the following LPM:

$$\widehat{inlf} = 0.052 + 0.046 \cdot educ - 0.224 \cdot kidslt6$$

1. **What is the predicted probability of labor force participation for a woman who has 12 years of education and 2 children under the age of 6 years old? 3 children under the age of 6 years old?**

2. **Interpret the coefficient on *educ*.**

3. **For a woman with 16 years of education, what is the predicted change in probability of labor force participation when going from 0 to 1 young child? From 1 young child to 2?**

- **Advantages/Disadvantages.**

- **Statistical Inference.**

- **Measures of Fit.**

2 Probit

- **Specification.**

- **Estimation Method.**

- **Example.** Consider the following probit results:

$$\widehat{P(inlf|educ, kidslt6)} = \Phi(-1.259 + 0.129 \cdot educ - 0.621 \cdot kidslt6)$$

1. **What is the predicted probability of labor force participation for a woman who has 12 years of education and 2 children under the age of 6 years old? 3 children under the age of 6 years old?**

2. **Interpret the coefficient on** *educ***.**

3. **For a woman with 16 years of education, what is the predicted change in probability of labor force participation when going from 0 to 1 young child? From 2 young children to 3?**

- **Advantages/Disadvantages.**

- **Statistical Inference.**

- **Measures of Fit.**

3 Logit

- **Specification.**

- **Estimation Method.**

- **Example.** Consider the following probit results:

$$\widehat{P(inlf|educ}, kidslt6) = \Lambda(-2.053 + 0.210 \cdot educ - 1.010 \cdot kidslt6)$$

1. **What is the predicted probability of labor force participation for a woman who has 12 years of education and 2 children under the age of 6 years old? 3 children under the age of 6 years old?**

2. **Interpret the coefficient on** *educ***.**

3. **For a woman with 16 years of education, what is the predicted change in probability of labor force participation when going from 0 to 1 young child? From 2 young children to 3?**

- **Advantages/Disadvantages.**

- **Statistical Inference.**

- **Measures of Fit.**

4 Exercises

Exercise 1. (Review the Concepts, Chapter 11.) Suppose that a linear probability model yields a predicted value of Y that is equal to 1.3. Explain why this is non-sensical.

Exercise 2. (Review the Concepts, Chapter 11.) On of your friends is using data on individuals to study the determinants of smoking at your university. She asks you whether she should use a probit, logit or linear probability model, what advice do you give here? Why?

Exercise 3. (Review the Concepts, Chapter 11.) Why are the coefficients of the probit and logit models estimated by maximum likelihood instead of OLS?

Exercise 4. You want to estimate how income affects the probability of voting for a republican candidate. The result from your logit model is the following:

$$\widehat{P(Republican = 1|Income)} = \Lambda(-1.00 + 0.02 Income)$$

where $Income$ is measured in thousands of dollars.

(a) What is the probability of voting republican if the income is 10,000 dollars?

(b) What is the marginal effect of income, evaluated at the sample mean of income of 50,000 dollars?

Final Exam 2009, Question 3. In order to better target their marketing, a cell phone company wants to assess the determinants of cell phone subscribership. Their consultant estimates the following Linear Probability Model (LPM) using a sample of individuals older than 18 years of age:

$$CELL_i = \underset{(0.121)}{0.30} - \underset{(0.005)}{0.02}\, AGE_i + \underset{(0.020)}{0.05}\, EDUC_i + \underset{(0.006)}{0.10}\, ln(INCOME_i)$$

where $CELL$ is a dummy variable equal to 1 if person i subscribes to a cell service, 0 otherwise.

(a) What is the interpretation of the estimate of the coefficient on ln(INCOME)?

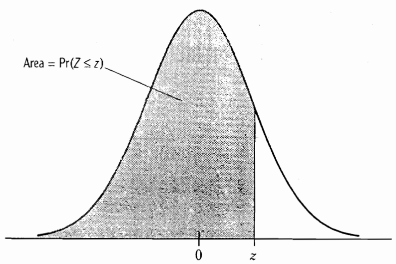
(b) Suppose you want to forecast the probability that a 50-year-old individual with 12 years of education, and income of $40,000 will own a cell phone. What prediction does this model make and does it make economic sense? [HINT: ln(40,000) = 10.6.]

(c) Describe two problems with least squares estimation of the coefficients in the LPM when some variables, such as ln(INCOME), can get very large or even infinite.

(d) Instead of using a LPM, you recommend to the consultant to estimate a probit model using the dummy dependent variable CELL. Give the specification of the model and explain how to compute the estimates of the coefficients in the probit model.

Appendix

TABLE 1 The Cumulative Standard Normal Distribution Function, $\Phi(z) = \Pr(Z \le z)$

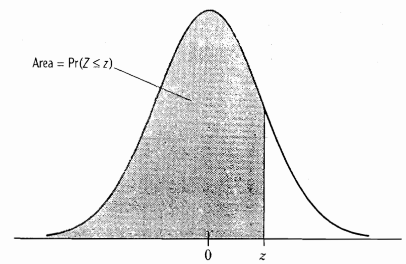


z	Second Decimal Value of z: 0	1	2	3	4	5	6	7	8	9
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611

(Table 1 continued)

Appendix

TABLE 1 The Cumulative Standard Normal Distribution Function, $\Phi(z) = \Pr(Z \le z)$



z	Second Decimal Value of z: 0	1	2	3	4	5	6	7	8	9
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611

(Table 1 continued)

749