MyEconLab Can Help You Get a Better Grade

MyEconLab If your exam were tomorrow, would you be ready? For each chapter, **MyEconLab** Practice Tests and Study Plan help you prepare for your exa
You can also find the Exercises and all Review the Concepts Questions available now in **MyEconl**
To see how it works, turn to the **MyEconLab** spread on the inside front cover of this book and th
go to **www.myeconlab.com**.

For additional Empirical Exercises and Data Sets, log on to the Companion Website at **www.pearsonhighered.com/stock_watson.**

Review the Concepts

11.1 Suppose that a linear probability model yields a predicted value of Y is equal to 1.3. Explain why this is nonsensical.

11.2 In Table 11.2 the estimated coefficient on *black* is 0.084 in colum 0.688 in column (2), and 0.389 in column (3). In spite of these large d ences, all three models yield similar estimates of the marginal effect of on the probability of mortgage denial. How can this be?

11.3 One of your friends is using data on individuals to study the dete nants of smoking at your university. She asks you whether she sh use a probit, logit, or linear probability model. What advice do you her? Why?

11.4 Why are the coefficients of probit and logit models estimated by maxi likelihood instead of OLS?

Exercises

Exercises 11.1 through 11.5 are based on the following scenario: Four hun driver's license applicants were randomly selected and asked whether they pa their driving test ($Pass_i = 1$) or failed their test ($Pass_i = 0$); data were also lected on their gender ($Male_i = 1$ if male and = 0 if female) and their yea driving experience ($Experience_i$, in years). The following table summarizes sev estimated models.

11.1 Using the results in column (1):

a. Does the probability of passing the test depend on *Experience*? Explain.

ependent Variable: *Pass*

	Probit (1)	Logit (2)	Linear Probability (3)	Probit (4)	Logit (5)	Linear Probability (6)	Probit (7)
perience	0.031 (0.009)	0.040 (0.016)	0.006 (0.002)				0.041 (0.156)
le				−0.333 (0.161)	−0.622 (0.303)	−0.071 (0.034)	−0.074 (0.259)
le × Experience							−0.015 (0.019)
stant	0.712 (0.126)	1.059 (0.221)	0.774 (0.034)	1.282 (0.124)	2.197 (0.242)	0.900 (0.022)	0.806 (0.200)

b. Matthew has 10 years of driving experience. What is the probability that he will pass the test?

c. Christopher is a new driver (zero years of experience). What is the probability that he will pass the test?

d. The sample included values of *Experience* between 0 and 40 years, and only four people in the sample had more than 30 years of driving experience. Jed is 95 years old and has been driving since he was 15. What is the model's prediction for the probability that Jed will pass the test? Do you think that this prediction is reliable? Why or why not?

11.2 **a.** Answer (a) through (c) from Exercise 11.1 using the results in column (2).

b. Sketch the predicted probabilities from the probit and logit in columns (1) and (2) for values of *Experience* between 0 and 60. Are the probit and logit models similar?

11.3 **a.** Answer (a) through (c) from Exercise 11.1 using the results in column (3).

b. Sketch the predicted probabilities from the probit and linear probability in columns (1) and (3) as a function of *Experience* for values of *Experience* between 0 and 60. Do you think that the linear probability is appropriate here? Why or why not?

11.4 Using the results in columns (4) through (6):

a. Compute the estimated probability of passing the test for men and for women.

b. Are the models in (4) through (6) different? Why or why not?

the probability that a worker was employed in April 2009, conditio on being employed in April 2008.

b. Regress *Employed* on *Age* and Age^2, using a linear probability mod

i. Based on this regression, was age a statistically significant deter nant of employment in April 2009?

ii. Is there evidence of a nonlinear effect of age on the probability being employed?

iii. Compute the predicted probability of employment for a 20-year old worker, a 40-year-old worker, and a 60-year-old worker.

c. Repeat (b) using a probit regression.

d. Repeat (b) using a logit regression.

e. Are there important differences in your answers to (b)–(d)? Explai

f. The data set includes variables measuring the workers' educational attainment, sex, race, marital status, region of the country, and weekly earnings in April 2008.

i. Construct a table like Table 11.2 to investigate whether the conclusions on the effect of age on employment from (b)–(d) are affected by omitted variable bias.

ii. Use the regressions in your table to discuss the characteristics of workers who were hurt most by the Great Recession.

g. The results in (a)–(f) were based on the probability of employment. Workers who are not employed can either be (i) unemployed or (ii) out the labor force. Do the conclusions you reached in (a)–(f) al hold for workers who became unemployed? (*Hint:* Use the binary variable *Unemployed* instead of *Employed*.)

h. These results have covered employment transitions during the Grea Recession, but what about transitions during normal times? On the textbook website, you will find the data file **Employment_06_07**, which measures the same variables but for the years 2006–2007. Ana lyze these data and comment on the differences in employment tran sitions during recessions and normal times.

E11.2 Believe it or not, workers used to be able to smoke inside office building Smoking bans were introduced in several areas during the 1990s. In addi tion to eliminating the externality of secondhand smoke, supporters o these bans argued that they would encourage smokers to quit by reducin their opportunities to smoke. In this assignment you will estimate the effe

of workplace smoking bans on smoking, using data on a sample of 10,000 U.S. indoor workers from 1991 to 1993, available on the textbook website, **http://www.pearsonhighered.com/stock_watson**, in the file **Smoking**. The data set contains information on whether individuals were or were not subject to a workplace smoking ban, whether the individuals smoked, and other individual characteristics.[7] A detailed description is given in **Smoking_Description**, available on the website.

a. Estimate the probability of smoking for (i) all workers, (ii) workers affected by workplace smoking bans, and (iii) workers not affected by workplace smoking bans.

b. What is the difference in the probability of smoking between workers affected by a workplace smoking ban and workers not affected by a workplace smoking ban? Use a linear probability model to determine whether this difference is statistically significant.

c. Estimate a linear probability model with *smoker* as the dependent variable and the following regressors: *smkban*, *female*, *age*, age^2, *hsdrop*, *hsgrad*, *colsome*, *colgrad*, *black*, and *hispanic*. Compare the estimated effect of a smoking ban from this regression with your answer from (b). Suggest a reason, based on the substance of this regression, explaining the change in the estimated effect of a smoking ban between (b) and (c).

d. Test the hypothesis that the coefficient on *smkban* is zero in the population version of the regression in (c) against the alternative that it is nonzero, at the 5% significance level.

e. Test the hypothesis that the probability of smoking does not depend on the level of education in the regression in (c). Does the probability of smoking increase or decrease with the level of education?

f. Repeat (c)–(e) using a probit model.

g. Repeat (c)–(e) using a logit model.

h. i. Mr. A is white, non-Hispanic, 20 years old, and a high school dropout. Using the probit regression and assuming that Mr. A is not subject to a workplace smoking ban, calculate the probability that Mr. A smokes. Carry out the calculation again, assuming that

[7]These data were provided by Professor William Evans of the University of Maryland and were used in his paper with Matthew Farrelly and Edward Montgomery, "Do Workplace Smoking Bans Reduce Smoking?" *American Economic Review*, 1999, 89(4): 728–747.

he is subject to a workplace smoking ban. What is the effect of the smoking ban on the probability of smoking?

ii. Repeat (i) for Ms. B, a female, black, 40-year-old college graduate.

iii. Repeat (i)–(ii) using the linear probability model.

iv. Repeat (i)–(ii) using the logit model.

v. Based on your answers to (i)–(iv), do the logit, probit, and linear probability models differ? If they do, which results make most sense? Are the estimated effects large in a real work sense?

APPENDIX

11.1 The Boston HMDA Data Set

The Boston HMDA data set was collected by researchers at the Federal Reserve Bank of Boston. The data set combines information from mortgage applications and a follow-up survey of the banks and other lending institutions that received these mortgage applications. The data pertain to mortgage applications made in 1990 in the greater Boston metropolitan area. The full data set has 2925 observations, consisting of all mortgage applications by blacks and Hispanics plus a random sample of mortgage applications by whites.

To narrow the scope of the analysis in this chapter, we use a subset of the data for single-family residences only (thereby excluding data on multifamily homes) and for black applicants and white applicants only (thereby excluding data on applicants from other minority groups). This leaves 2380 observations. Definitions of the variables used in this chapter are given in Table 11.1.

These data were graciously provided to us by Geoffrey Tootell of the Research Department of the Federal Reserve Bank of Boston. More information about this data set, along with the conclusions reached by the Federal Reserve Bank of Boston researchers, is available in the article by Alicia H. Munnell, Geoffrey M. B. Tootell, Lynne E. Browne, and James McEneaney, "Mortgage Lending in Boston: Interpreting HMDA Data," *American Economic Review*, 1996, pp. 25–53.

APPENDIX

11.2 Maximum Likelihood Estimation

This appendix provides a brief introduction to maximum likelihood estimation in the context of the binary response models discussed in this chapter. We start by deriving the MLE of the success probability p for n i.i.d. observations of a Bernoulli random variable. We then