

- c. What would be the effect on the probability of a nonwhite woman living on her own, if *Education* and *Family composition* were changed from their current mean to the mean of whites, while all other variables were left unchanged at the nonwhite mean values?
 - d. How would you calculate marginal effects in a more advanced statistical package? Why are marginal effects more informative than the coefficient estimates printed in this problem?
 - e. What are the primary reasons probit or logit models are used instead of the linear probability model?
- 12.4 Suppose you were hired by the state of California to conduct a study of voter willingness to support a tax increase to fund new school spending. Define V_i such that:
- $$V_i = 1 \text{ if voter } i \text{ supports the tax increase}$$
- $$V_i = 0 \text{ if voter } i \text{ does not support the tax increase}$$

Thus, V_i is a binomial random variable with population mean, μ , and population variance, σ^2 . You think that voter support of the tax increase is related to the income of each individual.

- a. Suppose that you regress V_i on Income_i using a linear probability model. What properties would this model have, assuming that income is the only factor that affects voting behavior? Is the estimate of the slope reliable?
- b. How would your answer change in part (a) if you instead used a logit or probit model?

Exercises

- E12.1 Use the file **Hunting Data.xls**, which is collected from the 2006 National Survey of Fishing, Hunting, and Wildlife Associated Recreation, for the following questions.
- a. Use OLS to regress *hunt* on *male*, *married*, *divorced*, *somcol*, *hsonly*, *ba*, *somepost*, *postgrad*, *black*, *other*, *income*, and *south central*. Comment on the results. Why is OLS not the appropriate estimation strategy in this circumstance?
 - b. Estimate the same model as in part (a) with a probit model. Comment on the results. Make sure to estimate the marginal effects.
 - c. Estimate the same model as in part (a) with a logit model. Comment on the results. Make sure to estimate the marginal effects.
 - d. Which model do you think is the most appropriate to use and why?
- E12.2 Using the file **Cigarette Data.xls**, which is collected from the 2003 National Health Interview Survey,
- a. Using OLS, regress *smoke* on *black*, *Hispanic*, *male*, *age*, *family size*, *married*, and *divorced*. Comment on the results.
 - b. Estimate the same model as in part (a) with a probit model. Comment on the results. Make sure to estimate the marginal effects.
 - c. Estimate the same model as in part (a) with a logit model. Comment on the results. Make sure to estimate the marginal effects.
 - d. Which model do you think is the most appropriate to use and why?
- E12.3 Using the file **Publication.xls**, estimate a multinomial logit to investigate the factors that lead a researcher to either not publish (*publish* = 0), be a moderate publisher (*publish* = 1), or be a prolific publisher (*publish* = 2). Make sure to obtain the marginal effects and comment on your results.
- E12.4 Using the file **Publication.xls**, estimate an ordered probit to investigate the factors that lead a researcher to either not publish (*publish* = 0), be a moderate publisher (*publish* = 1), or be a prolific publisher (*publish* = 2). Make sure to obtain the marginal effects and comment on your results.

Handwritten note:
 Use the file **Hunting Data.xls**
 Use the file **Cigarette Data.xls**
 Use the file **Publication.xls**