1. Interpretação do Modelo

   **C5**  In Problem 2 in Chapter 4, we added the return on the firm's stock, *ros*, to a model explaining CEO salary; *ros* turned out to be insignificant. Now, define a dummy variable, *rosneg*, which is equal to one if $ros < 0$ and equal to zero if $ros \geq 0$. Use CEOSAL1.RAW to estimate the model

   $$\log(salary) = \beta_0 + \beta_1\log(sales) + \beta_2 roe + \beta_3 rosneg + u.$$

   Discuss the interpretation and statistical significance of $\hat{\beta}_3$.

   **C8**  Use the data in LOANAPP.RAW for this exercise. The binary variable to be explained is *approve*, which is equal to one if a mortgage loan to an individual was approved. The key explanatory variable is *white*, a dummy variable equal to one if the applicant was white. The other applicants in the data set are black and Hispanic.

   To test for discrimination in the mortgage loan market, a linear probability model can be used:

   $$approve = \beta_0 + \beta_1 white + other\ factors.$$

   (i) If there is discrimination against minorities, and the appropriate factors have been controlled for, what is the sign of $\beta_1$?

   (ii) Regress *approve* on *white* and report the results in the usual form. Interpret the coefficient on *white*. Is it statistically significant? Is it practically large?

   (iii) As controls, add the variables *hrat*, *obrat*, *loanprc*, *unem*, *male*, *married*, *dep*, *sch*, *cosign*, *chist*, *pubrec*, *mortlat1*, *mortlat2*, and *vr*. What happens to the coefficient on *white*? Is there still evidence of discrimination against nonwhites?

   (iv) Now, allow the effect of race to interact with the variable measuring other obligations as a percentage of income (*obrat*). Is the interaction term significant?

2. Problemas de Especificação e Variável Proxy

   *Nota: Problema C1 se refere ao problema C5 acima*

   **C1**  (i) Apply RESET from equation (9.3) to the model estimated in Computer Exercise C5 in Chapter 7. Is there evidence of functional form misspecification in the equation?

   (ii) Compute a heteroskedasticity-robust form of RESET. Does your conclusion from part (i) change?

   **C2**  Use the data set WAGE2.RAW for this exercise.

   (i) Use the variable *KWW* (the "knowledge of the world of work" test score) as a proxy for ability in place of *IQ* in Example 9.3. What is the estimated return to education in this case?

   (ii) Now, use *IQ* and *KWW* together as proxy variables. What happens to the estimated return to education?

   (iii) In part (ii), are *IQ* and *KWW* individually significant? Are they jointly significant?

## 3. OLS e PROBIT

**C1**  Use the data in PNTSPRD.RAW for this exercise.

(i)  The variable *favwin* is a binary variable if the team favored by the Las Vegas point spread wins. A linear probability model to estimate the probability that the favored team wins is

$$P(favwin = 1|spread) = \beta_0 + \beta_1 spread.$$

Explain why, if the spread incorporates all relevant information, we expect $\beta_0 = .5$.

(ii)  Estimate the model from part (i) by OLS. Test $H_0: \beta_0 = .5$ against a two-sided alternative. Use both the usual and heteroskedasticity-robust standard errors.

(iii)  Is *spread* statistically significant? What is the estimated probability that the favored team wins when *spread* $= 10$?

(iv)  Now, estimate a probit model for $P(favwin = 1|spread)$. Interpret and test the null hypothesis that the intercept is zero. [*Hint*: Remember that $\Phi(0) = .5$.]

(v)  Use the probit model to estimate the probability that the favored team wins when *spread* $= 10$. Compare this with the LPM estimate from part (iii).

(vi)  Add the variables *favhome*, *fav25*, and *und25* to the probit model and test joint significance of these variables using the likelihood ratio test. (How many *df* are in the chi-square distribution?) Interpret this result, focusing on the question of whether the spread incorporates all observable information prior to a game.

## 4. PROBIT e LOGIT

**C2**  Use the data in LOANAPP.RAW for this exercise; see also Computer Exercise C8 in Chapter 7.

(i)  Estimate a probit model of *approve* on *white*. Find the estimated probability of loan approval for both whites and nonwhites. How do these compare with the linear probability estimates?

(ii)  Now, add the variables *hrat*, *obrat*, *loanprc*, *unem*, *male*, *married*, *dep*, *sch*, *cosign*, *chist*, *pubrec*, *mortlat1*, *mortlat2*, and *vr* to the probit model. Is there statistically significant evidence of discrimination against nonwhites?

(iii)  Estimate the model from part (ii) by logit. Compare the coefficient on *white* to the probit estimate.

(iv)  Use equation (17.17) to estimate the sizes of the discrimination effects for probit and logit.