

**COLORADO STATE UNIVERSITY
DEPARTMENT OF AGRICULTURAL & RESOURCE ECONOMICS**

*Problem Set 4
Fall 2023*

**Agricultural & Resource Economics / Economics 535
Applied Econometrics**

S.R. Koontz

Probit and Logit Models

Provide answers to the following questions in a word-processed document. This problem set is to be a team effort between two individuals. Each team must submit a unique assignment. However, you may work with other students. Show the work necessary to communicate clear answers each question. Good answers always communicate. This problem set is worth 50 points. Round all reported statistics to the fourth decimal

The objective of this problem set is to estimate qualitative choice models. You will need to estimate models and conduct hypotheses tests using econometric computer software. You will also need to interpret the results. The problem set is organized around modelling the vote behavior of individuals on a local school millage (i.e., tax assessment) referendum in Troy, Michigan. The model examines if socio-economic variables explain voter choice. The data you will need are on the course website.

A description of the data is as follows: Case is an observation number counter. YES is 1 if voter favored the millage referendum. P1&2 is 1 if voter has one or two children in public school. P3&4 is 1 if voter has three or four children in public school. P5 is 1 if voter has five or more children in public school. PRI is 1 if voter has one or more children in private school. YRS equals the number of years the voter has lived in the Troy community. SCH is 1 if voter is employed by the public school system. INCOME equal annual income of voter's household. PROPTX equals annual property taxes paid by the voter's household. YES, P1&2, P3&4, P5, PRI, and SCH are dummy variables. YRS, INCOME, and PROPTX are continuous variables.

Remember to practice good data analysis habits. Familiarize yourself with the data: print and visually examine the data, and also examine histograms of each variable. Answer the following questions.

1. Report a table of summary statistics and a table of correlations.
2. Estimate and report OLS results for the following model

$$\text{YES} = f(\text{PUB1\&2}, \text{PUB3\&4}, \text{PUB5}, \text{PRI}, \text{YRS}, \text{SCH}, \ln(\text{INCOME}), \ln(\text{PROPTX})).$$

This is the linear probability model. Present your results in a table. The table should include parameter estimates, standard errors, p-values, and summary statistics for the model. Because the model has heteroskedasticity, also report White's heteroskedasticity consistent standard errors and p-values. Does heteroskedasticity appear to be an important problem with this sample? Examine the predicted values from the model. Are there any predictions less than zero or greater than one?

3. Construct some contingency tables between YES and the other dummy variables that are in the model. Graph YES versus $\ln(\text{INCOME})$ and YES versus $\ln(\text{PROPTX})$. Does the graph communicate?

4. Estimate and report Probit and Logit results for the following the same model

$$\text{YES} = f(\text{PUB1\&2}, \text{PUB3\&4}, \text{PUB5}, \text{PRI}, \text{YRS}, \text{SCH}, \ln(\text{INCOME}), \ln(\text{PROPTX})).$$

Present your results in a table. The table should include parameter estimates, standard errors, p-values, summary statistics for the model, and the change in the probability of voting for the referendum given a change in each independent variable. Discuss the statistical significance of the coefficients and of the overall model. Calculate and interpret the percent correct classifications. Discuss the economic interpretation of the estimates.

5. The result associated with the YRS variable appears quite small. But it is the one variable that is not a dummy variable and nor is it logarithmic. Figure out how to calculate an elasticity without re-estimating the models, using with the impact probability, and report it.
6. Calculate the probability that each voter favored the referendum given the voter had the following sets of attributes.
 - a. Two children in public school, no children in private school, does not work for the public school system, 10 years in the community, household income of \$20,000, and property taxes of \$1,150.
 - b. Three children in public school, no children in private school, works for the public school system, 5 years in the community, household income of \$17,000, and property taxes of \$1,050.
 - c. No children in public school, one or more children in private school, does not work for the public school system, 20 years in the community, household income of \$14,500, and property taxes of \$1,250.
7. Compare the results of the Linear Probability Model, Probit Model and Logit Model. Are the results qualitatively different? Discuss.