COLORADO STATE UNIVERSITY
DEPARTMENT OF AGRICULTURAL & RESOURCE ECONOMICS

Problem Set 2 **Agricultural & Resource Economics / Economics 535** *S.R. Koontz*
Fall 2023 **Applied Econometrics**

Estimating Regression Equations and Hypothesis Testing

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 discuss your work with other students. Show the work necessary to communicate clear answers to each question. Good answers always communicate. This problem set is worth 60 points. Round all reported statistics to the fourth decimal.

The objective of this problem set is to continue conducting applied regression analysis. You will need to estimate regressions and conduct specific hypotheses tests using statistical software. The problem set is organized around estimating a demand function for the commodity beef. The data you will need are on the class website. You have been provided with data on per capita consumption of beef, pork and chicken, real retail-level prices for beef, pork and chicken, and real disposable personal income from 1960 through 2022. Prices and income are deflated to \$2008. Remember to practice good data analysis habits. Familiarize yourself with the data: examine plots of the data against time and examine pair-wise plots of the variables. Answer the following questions.

1. Report a table of summary statistics and a table of correlations for the variables used in the regression model below. Report summary statistics of the data in levels and correlations of the data in natural logarithms. Construct a plot the real beef price against per capita consumption. Label the year for each point.

2. Estimate and report the beef demand equation regression of the following form

$$\ln(y_t) = \beta_0 + \beta_1 \ln(x_{1t}) + \beta_2 \ln(x_{2t}) + \beta_3 \ln(x_{3t}) + \beta_4 \ln(x_{4t}) + e_t$$

where y = per capita consumption of beef, lbs.
x_1 = real beef price, cents per lb.
x_2 = real pork price, cents per lb.
x_3 = real chicken price, cents per lb.
x_4 = real disposable personal income, dollars.

This model is usually referred to as a double-log model indicating that variables on the left-hand and right-hand sides are both converted to natural logarithms. Write the estimates in equation form. (See equation 5.11.1 on page 129 in Gujarati for an example.) Interpret the slope coefficients. The coefficients are a particular economic measure. Use that measure in your interpretation. (See Gujarati pages 159-162.) Discuss the statistical significance of the coefficients and of the overall regression. Interpret the model R^2.

3. Does the economic interpretation of the coefficient estimates agree with *a prior* expectations based on consumer theory? Be concise.

4. Test equality of the coefficients on the other-meat price variables. This test addresses the question of whether or not substitute meat prices have the same relative impact on beef

consumption. Write down your null hypothesis, report the test statistic, the distribution of the test under the null hypothesis, and p-value. State the conclusion of the test.

5. Perform a joint test of the significance of the other-meat price variables and the income variable. This test addresses the question of whether or not demand shifting variables impact on beef consumption. Write down your null hypothesis, report the test statistic and p-value. State the conclusion of the test.

The following two questions require you to use the results of the regression in some basic calculations.

6. Using the prices and income for 1986, calculate the expected value of beef consumption measured in pounds per capita. Using the prices and income for 1993, calculate the expected value of beef consumption measured in pounds per capita. (Don't forget to anti-log the expected value of consumption.) How much of the change in the expected value of beef consumption between the two years is due to the own-price effect or is a movement along the demand curve? (Be careful here. Read Gujarati pages 159-162.)

7. Given the other-meat prices, income, and per capita beef consumption in 1998, solve for the real beef price. Given the other-meat prices, income, and per capita beef consumption in 2010, solve for the real beef price. The prices should be measured in cents per pound.

The regression results for this demand model suggest the model is problematic. The results do not agree with what you would expect from consumer theory. There are several potential reasons why. You will examine two. The first is that the independent variables are collinear. OLS cannot sort out the effects of each independent variable on the dependent variable, *ceteris paribus*. One method of addressing this problem is to incorporate restrictions on the coefficients.

8. Test whether or not the demand model supports homogeneity. Homogeneity implies the sum of the own-price elasticity, cross price elasticities, and income elasticity equal zero. Test whether or not the data support the restriction. Testing a restriction requires that it not be enforced. This is the unrestricted model you estimated earlier. You do not need to report the unrestricted model again. Just conduct and report the test.

9. Estimate regression enforcing homogeneity. This is the restricted model. Estimate and report this model. Your results must have all the elasticities, standard errors, and p-values.

The second reason why there may be problems with the model is that there is much evidence that the structure of meat demand changed over the sample period. What OLS assumption does this violate? (This is a rhetorical question.) Examine the plot of real beef prices and per capita consumption. What do you see happening through the 1980s and 1990s? (Another rhetorical question.)

10. Test for a structural change in the demand function during and after 1981. Assume the change occurs in 1981 and persists all years thereafter. Use dummy variables in a Chow Test based on the dummy variables. (See Gujarati pages 254-259 and 285-288.) Is there evidence of structural change? (This is not a rhetorical question – perform a statistical test.) What coefficients changed the most in the later period? Be concise.

11. A Chow Test is relatively easy to perform but assumes that the error variance is the same before and after the potential structural change. This is potentially a limiting assumption. Test whether or the error variance is the same before and after the assumed structural change point in time. (See Gujarati page 258.) Preform a one-tailed test (i.e., put the larger variance on the top of the

ratio) and report the test result and p-value. (Regardless of the outcome stop here with the Chow Test structural change testing.)

12. As opposed to a one-time change in the demand function, the structural change could be due to a gradual change in consumer tastes and preferences. Incorporate three trend variables into the model, re-estimate the regression, and report the results. The first trend variable should be the observation number, the second should be the observation number squared, and the third should be the observation number cubed. (Be sure to read Gujarati page 164 for a discussion of trend variables and pages 210-213 for a discussion of polynomial variables.) Are the trend variables jointly significant?

13. Describe the trend in beef consumption. You could take the derivative of the dependent variable with respect to the trend variable. This will be an equation which is a function of the year variable. You could plug the sequence of years into this equation and calculate the change in the expected value of consumption. But you would have to worry about the mean correction. An easier way is to plug the mean value of prices and income into your model and then plug in the sequence of years. This way you are holding the prices and income constant at their mean value and allowing consumption to change based only on the trend. Graph the expected value of consumption over time. The result is the change in consumption that year due to changing tastes and preferences (or at least that's what we are assuming the trend measures). What years does the trend change directions?

This problem set is more than a fair bit of work but here is clearly a model building process. The logical next step is to combine the dummy variables and the trend variables and see if the encompassing model rules out either or says both types of structural change should be included. There is also a potential second structural change after 1999 requiring another set of dummy variables. And we could do all while testing for and potentially enforcing homogeneity. Please do not do this.

14. Graph the per capita consumption data. Graph the predicted values from the OLS model, the OLS model with the homogeneity restriction, and the OLS model with the trend values. Which model follows the data better?

Optional Questions

O1. Conduct collinearity diagnostics on the independent variables used in the double-log model. Is there problematic collinearity in the data? Collinearity between the trend variables is less important because we use them in combination.

O2. Conduct analysis to look for influential points in the data. Are there outliers? What are the years? What β's are impacted the most?

O3. An alternative to the double-log model is a model which is linear in variables. This model is written as

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + e_t$$

where y and x variables are as defined earlier. Conduct a test and determine which functional form is more appropriate. (See Gujarati pages 260-261.) One of the reasons that the double-log model may have so many problems is that it may be the wrong functional form for the data. We use the double-log model because it generates elasticities as parameters. But this is not so good if the functional form does not fit the data. What do you think?