COLORADO STATE UNIVERSITY
DEPARTMENT OF AGRICULTURAL & RESOURCE ECONOMICS

Problem Set 1 **Agricultural & Resource Economics / Economics 535** *S.R. Koontz*
Fall 2023 **Applied Econometrics**

Estimating, Interpreting, and Using Regression Equations

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 each question. Good answers always communicate. This problem set is worth 40 points. Round all reported statistics to the fourth decimal.

The objective of this problem set is to introduce you to the process of conducting applied regression analysis. Students in the course have a varying depth of skill with statistical software packages. Those with experience estimating regressions will complete this work quickly. Those students with less experience are encouraged to use this assignment as motivation to begin learning an econometric software package. Using MS Excel for graphics is acceptable but for the regression analysis is not acceptable.

Problem 3.22 in Gujarati and Porter (page 88) is the basis of this assignment. This is the classic problem of what asset is a better hedge against inflation: stocks or gold? The data have been updated and includes observations from 1970 through 2022. Instead of the NY Stock Exchange index the Willshire 5000 Index (WILL5000IND) is used for stock prices. The London gold price is used and is quoted in US dollars (from NASDAQ). The ICE BofA Corporate Bond Total Market Index (BAMLCC0A0CMTRIV) is also included as an investment. So, we are asking which asset is a better hedge against inflation: stocks, bonds, or gold?

1. Construct seven data plots. Plot the Stock Index, Bond Index, Gold Price, and CPI against the year on separate plots. Next, plot the data used in the regression models. Plot the Stock Index against CPI, the Bond Index against CPI, and Gold Price against the CPI where all the data are natural logarithms.

Plotting your data is simply a habit. Examining plots will help the researcher develop an understanding of the variation in each series. The process also helps the researcher identify errors in data entry. There is no excuse for data errors.

2. Report a table of summary statistics for the Stock Index, Bond Index, Gold Price, and CPI series. Summary statistics should include the mean, standard deviation, minimum, maximum, and number of observations. Report the data in levels and logarithms. Also report a table of correlation coefficients between the series used in the regression models. Comment.

Reporting your sample period and summary statistics is also a good habit to develop. The process helps the researcher communicate knowledge of data characteristics to the reader. It is also a must if the reader wants to work with model results such as calculate expected values or elasticities.

3. The price of potential assets should parallel measures of inflation to be useful hedges against inflation. Stocks, bonds, and gold are three potential assets. To assess which asset price moves with inflation, estimate the following three models using OLS.

$$\ln(\text{Stock Index}_t) = \beta_0 + \beta_1 \ln(\text{CPI}_t) + e_t$$

$$\ln(\text{Bond Index}_t) = \beta_0 + \beta_1 \ln(\text{CPI}_t) + e_t$$

$$\ln(\text{Gold Price}_t) = \beta_0 + \beta_1 \ln(\text{CPI}_t) + e_t$$

Write the estimates in equation form. Report and identify the standard error, t-statistic, and p-value of each coefficient in parenthesis under the estimate. (See equation 5.11.1 on page 129 in Gujarati for a formatting example.) Also report the R-square, F-statistic, and square root of the error variance estimate σ some programs call this the standard error of the regression (error).

4. Provide an exact interpretation of the slope coefficient in the Stock Index model. Do the same for the other models.

5. If the CPI = 130.0 in 2023, what is the expected value of the Stock Index? What is the expected value of the Bond index? What is the expected value of the Gold Price?

6. Your regression package should report t-statistics for each coefficient, an F-statistic for each regression, and p-values for each statistic. Concisely, show the calculations for each slope coefficient t-statistic. Interpret the p-values for the slope coefficient t-statistics and the regression F-statistics.

7. Test the null hypothesis that the slope coefficient from each model equals one (H_0: $\beta_1 = 1$). Report the three t-tests and p-values for each test. Interpret each p-value.

8. Which asset will best protect the value of your portfolio from inflation? Justify your choice based only on your statistical results. (And remember to take the specification as given.)

9. Next, you will need to re-scale the logarithmic CPI data. Subtract the mean from each observation and divide each observation by 10. This re-scaling "centers" the CPI data on its mean and converts the CPI data from logged points to tens of logged points. Re-estimate the models with OLS using the re-scaled CPI data. Write the estimates in equation form. Report the standard error, t-statistic, and p-value of each coefficient in parenthesis under each estimate. Report the model R-square, square root of the error variance estimate, and F-statistic. Examine the slope and intercept coefficients for each model. Have we fundamentally changed the models? Or are OLS models invariant to linear re-scaling of the data?

10. You are using data in logged "levels" in the models so far. Next, I want you to use data in "changes."

$$\Delta x_t = x_t - x_{t-1}$$

However, changes can be more useful when compared to a base. Therefore, use percent changes. Percent changes are simple to construct

$$\%\Delta\, x_t = (x_t - x_{t-1}) / x_{t-1} \approx (\ln(x_t) - \ln(x_{t-1})) = \Delta(\ln(x_t))$$

and because your data are annual you have calculated the year-to-year percent changes in the Stock Index, Bond Index, Gold Price, and CPI. Use the approximation and not the actual percent change. Report a table of summary statistics for the Stock Index, Bond Index, Gold Price, and

CPI percent changes series. Summary statistics should include the mean, standard deviation, minimum, maximum, and number of observations. Also report a table of correlation coefficients between the series. Comment.

11. Use the percent changes data to estimate your three models using OLS. The price of potential assets should parallel measures of inflation to be useful investments. Again, stocks, bonds, and gold are potential assets.

$$\%\Delta\text{Stock Index}_t = \beta_0 + \beta_1 \%\Delta\text{CPI}_t + e_t$$

$$\%\Delta\text{Bond Index}_t = \beta_0 + \beta_1 \%\Delta\text{CPI}_t + e_t$$

$$\%\Delta\text{Gold Price}_t = \beta_0 + \beta_1 \%\Delta\text{CPI}_t + e_t$$

Write the estimates in equation form. Report and identify the standard error, t-statistic, and p-value of each coefficient in parenthesis under the estimate. (See equation 5.11.1 in Gujarati for an example.) Also report the R-square, F-statistic, and square root of the error variance estimate.

12. Provide an exact interpretation of the slope coefficient in the Stock Index model. Do the same for the other models.

13. If an asset price mirrored inflation then the intercept in each model would equal zero and the slope coefficient would equal one. For each model, test the joint null hypothesis that the intercept equals zero and the slope coefficient equals one (H_0: $\beta_0 = 0$ & $\beta_1 = 1$). Report the test and p-values for each test. Interpret the p-value.

That was easy if you remember the percent changes part. Here's the trick question that requires some thought.

14. Which asset will best protect the value of your portfolio from inflation? Justify your choice based on your statistical results only. Is your answer different from number 8 above? If it is then you need to explain why. Think about the fundamental difference between the data series used in the two sets of models. Moving from levels to changes is not a linear re-scaling of the data. (And constructing a percentage is not the issue.) Can your advice from the two sets of models be designed for different types of investors – investors with different investment horizons?

Optional

For each of the three assets estimate the following model

$$\Delta(y_t) = \beta_0 + \beta_1 \Delta(x_t) + \beta_2 (y_{t-1} - x_{t-1}) + e_t$$

where the independent and dependent variables are as above – each is a natural logarithm of an asset price or the CPI. Notice that $z_t = z_{t-1} + \Delta z_t$. Plug this definition into the first model specification and derive the optional model specification. This optional model is a combination of the short-run and long-run models. But wait! Instead of $(y_{t-1} - x_{t-1})$ in the specification above use the appropriate residual from the first batch of models. The residual will account for a mean difference in the two variables. Briefly, discuss the implications of the results from the two-slope parameter estimates in each model.