**AGR 333 Data Science for Agriculture**

**Agricultural Economics Assignment**

For this assignment, you will draw on the exercises you performed in Labs 14 and 15. While most of the information you need can be found in the labs, some steps may require additional research. You will be asked to provide visuals (graphs and tables) and write a brief response about the underlying economics. **Upload your Lab 14 and Lab 15 code to GitHub and include the repository link in your submission.**

**Lab 14 Questions** (WASDE.csv)

**GitHub repository link to the code for Lab 14 (**2.50 points**):**

**0.50 points each question**

1. Make a scatter plot of corn prices on the y-axis and the stock-to-use ratio (SUR) on the x-axis with a linear regression line. What is the general relationship between corn prices and the SUR? Is a linear regression line a good representation of this relationship? A graph with a number of dots

   Description automatically generated

Generally, it’s ok. There is clear correlation between the two variables, and the points trend to regress towards the average near the end, however in early years there were clear outliers, and the trend line seems to be linear while the decrease over time continues to slow down. So, while the line does show the general direction of the change and correlation, it doesn’t accurately represent the change in pace of decreases over time as early the stock-to-use ratio really plummeted. Though they also had outliers on the negative side so it works both ways.

1. Run the linear regression: and show the regression output in a table. Include the regression coefficients, their standard errors, and the R-squared value. How do you interpret the coefficient on SUR? Is this consistent with the theory of supply and demand? How do you interpret the R-squared value? Describe the overall fit of this model.A screenshot of a table

   Description automatically generated

The R squared value is really low. Seems as if when corn prices go down, so does the stock price. However we can’t really be sure because the model isn’t really a good fit considering the extremely low R squared value. Though it seems right because when the production of corn goes up making the supply really large and also pushing the price down, the stock price should go down.

1. Run the non-linear regression: and show the regression output in a table. Include the regression coefficients, their standard errors, and the R-squared value. From this model, what is the marginal effect of SUR on the price of corn? Hint: if then .

A screenshot of a graph

Description automatically generated

We can just assume that as SUR increases, that price of corn decreases.

1. Create a scatter plot of corn prices on SUR with regression lines for two distinct periods: 1973-2005 vs. 2006-2019. How did the relationship change between the two time periods? What could explain this?

A graph of a graph with red and blue lines

Description automatically generated

In later years, rate of decreasing is much faster, but also the prices and scale are much larger. I think the simple explanation for this is just inflation, and the scale of goods in general is just much higher and increases at a larger rate each year considering the proportions of money.

1. Run the linear regression with a dummy variable for the post 2006 period and an interaction between the post-2006 dummy and SUR: and show the regression output in a table. Include the regression coefficients, their standard errors, and the R-squared value. From this model, what is the marginal effect of SUR on the price of corn in the 1973-2005 period? What is the effect of SUR in the 2006-2019 period? How do you interpret the coefficient on the P2006 dummy variable?

Decreases still in the 1973-2005 range, however it’s much smaller, the decrease is much higher in the later time range. Though this model is not good fit as the R squared value is really low

A screenshot of a table

Description automatically generated

**Lab 15 Questions** (soybeans-prices-monthly.csv, CPI.csv)

**GitHub repository link to the code for Lab 14 (**2.50 points**).**

**0.42 points each question**

1. Graph the trends in the nominal (not inflation adjusted) and real (inflation adjusted) price of soybeans. These can be two separate graphs or one plot with both series. What explains the difference between the two trends? Are these series stationary? Why or why not?

A graph of red and blue lines

Description automatically generated

Nominal Price goes up over time, and real price goes down. Soybean prices are going up, but that doesn’t mean they’re more valuable, likely due to inflation. In fact, they’re becoming less important. Not stationary because they are clearly changing. As pre 1990 the real price was always above the nominal, now it’s the complete opposite.

1. Decompose the nominal soybeans prices time-series and show the seasonal effects in a table by month. Which months have the largest seasonal effects?

0.10475698 -0.07697355 0.07815836 0.15075519 0.27893004

[6] 0.35801805 0.19607616 -0.08604261 -0.16743134 -0.28141584

[11] -0.17160982 -0.17370766

June has the biggest positive change, and October the biggest negative change

1. Graph the simple moving average (SMA) of nominal prices for 3, 12, and 48 months. Describe what happens as you increase the “order” (number of months) included in the SMA.

A graph of different types of prices

Description automatically generated with medium confidence

The lines become less sharp and rounded.

1. Graph the log differenced time-series of prices (i.e. percent returns). Does this series appear stationary?A graph of a graph showing the price of a change

   Description automatically generated with medium confidence

It does seem stationary. While there’s obviously a lot of change, it bounces around a single center point, I would guess that the average price change over the course is 0 or close to it.

1. Run an auto-regressive model on log-differenced prices and report the coefficients associated with each lag. The number of coefficients you report will depend on the number of lags you choose to put in your model. I.e., if you run an AR(5), you will report five coefficients. You do not need to report standard errors or R-squared, just the coefficients.

Coefficients:

1 2 3

0.3619 -0.0326 -0.1180

Order selected 3 sigma^2 estimated as 0.003398

1. Using your AR model, forecast future percent returns for the six months after the end of the time series (i.e. Aug, 2020-Jan, 2021). Graph your forecast along with the most recent 36 months. How confident are you in your forecast? How close was your forecast to the actual percent returns in soybeans prices for July-Oct of 2020?

A graph of a graph

Description automatically generated with medium confidence

Not confident at all because the confidence intervals are wide so you don’t really know a specific number that it will be. You can be confident but that it’s in such a wide range so it doesn’t really help at all.