Ecn 102 - Analysis of Economic Data
University of California - Davis
John Parman

February 24, 2011

Problem Set 5

This problem set will be due by 5pm on Friday, March 11th in your TA's mailbox in the economics department mailroom. You may work in groups but everyone in the group must write up their own solutions including creating their own graphs and tables. Include any relevant regression results, calculations and graphs from Excel with your solutions but do not include the raw data.

Analyzing Energy Consumption

In this problem set, you will use multivariate regressions to analyze energy consumption by American households. The data for this problem set are contained in the file *energy-use.csv* in the data folder on Smartsite. These data are a subset of the Residential Energy Consumption Survey available through the www.data.gov website. The file on Smartsite contains data on single-family detached houses and has the following set of variables:

- **hd65** - heating degree-days (this is a measure of how much heating is required over the year to warm the house up to 65 degrees, a decrease in the outside temperature of one degree for one day of the year would increase hd65 by one unit)
- **cd65** - cooling degree-days (this is a measure of how much cooling is requried over the year to cool the house down to 65 degrees, an increase in the outside temperature of one degree for one day of the year would increase cd65 by one unit)
- **totrooms** - number of rooms in the house
- **yearbuilt** - year in which the house was built
- **washload** - number of loads of laundry done each week (this is not the exact definition of this variable in the survey but it will work for our purposes)
- **kwh** - kilowatt-hours of electricity used annually
- **solar** - solar power dummy (equals one if household uses solar power for any purpose, equals zero if household does not use solar power)

a. Most households use air conditioning powered by electricity to cool down the house but use other forms of energy (gas, oil, etc.) to warm up the house. Given this piece of information, what would you predict for the sign and significance of the coefficients if electicity usage were regressed on a household's heating requirements and cooling requirements? Run a regression of electricity usage (kwh) on heating degree-days (hd65) and cooling degree-days (cd65). Are your results consistent with your predictions?

b. We'll focus on the relationship between cooling degree-days and electricity usage for the remainder of the problem. We could just run a simple regression of electricity usage (kwh) on cooling degree-days (cd65). However, we have additional variables that we could include that will affect electricity usage and would improve the explanatory power of our regression. Run a regression of electricty usage (kwh) on cooling degree-days (cd65), washer loads per week (washload), rooms in the house (totrooms) and whether the house uses solar power (solar). Are all of the signs of the coefficients what you would expect? Explain your answer for each coefficient.

c. An average washing machine uses a little less than one kilowatt-hour of electricity per load. Assuming that one load of laundry requires one kilowatt-hour of electricity, what would you expect the magnitude of the coefficient on washer loads per week to be? How does this compare to the coefficient you got for the number of washer loads? Give a possible explanation for any discrepencies between the size of your actual coefficient and the magnitude you expected based on the electricity use of an average washing machine. (Hint: Think about other variables that affect electricity usage and are not included in our regression but are correlated with the number of loads of laundry).

d. Would you expect the amount of electricity generated annually by a house's solar panels to be less than, equal to or greater than the value of the estimated coefficient for the solar dummy? Explain your answer.

e. It takes different amounts of energy to cool different houses down by one degree. For example, a large house will require more energy to cool down by a degree because a greater volume of air needs to be cooled. We can capture these effects through interaction terms. Create an interaction term between the number of rooms and cooling degree-days and another interaction term between the number of washer loads and the cooling degree-days. Rerurn your regression from part (b) including these interaction terms. Are the signs and significance of the coefficients on the interaction terms what you would expect? Why or why not? (You can assume that washer loads are a proxy for family size. In other words, the important change captured by an increase in washer loads is an increase in family size.)

f. Use an F test to determine whether including these interaction terms improved the regression. In other words, test whether the coefficients for the interaction terms are both zero or whether at least one of the coefficients is different from zero.

g. Based on your regression results from part (e), draw a line graph showing the predicted value of annual electricity usage as a function of the number of cooling degree-days for a household with five rooms that does one load of laundry a week and does not use solar power. Label the slope and intercept of your line with the appropriate values. On the same graph, draw another line showing the predicted value of annual electricity usage as a function of the number of cooling degree-days for a household with ten rooms that does one load of laundry a week and does not use solar power. Label the slope and intercept. You can draw this graph by hand. (It is not important that your graph is drawn to scale. What is important is that you label the slopes and intercepts correctly.)