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

January 22, 2011

Problem Set 3 - Solutions

This problem set will be due by 5pm on Monday, February 7th. It may be turned in either in class or put 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. Please only turn in what the questions asks for (do not print out and turn in the tables of raw data or variables you construct).

---

1. **Fuel Economy and Car Size** For this question, use the data on fuel economy for 2009 model cars. These data are contained in the file '2009-fuel-economy.csv' on Smartsite in the data folder. Descriptions of the variables are given in the file 'fuel-economy-variables.csv'.

   (a) The EPA's fuel economy data distinguishes between two-door, four-door and hatchback vehicles when specifying passenger and cargo space. As a result, in the data's current form, different variables capture car size for different car types. For our analysis, we will focus on four-door vehicles. Sort the data by the four-door passenger space variable and keep only those observations with a positive value reported for this variable (deleting the observations for two-door cars and hatchbacks). Create a new variable the measures that combined passenger and cargo space for the four-door cars.

   See *ps-3-solutions.xlsx*. Note that in the Excel solutions data, I have deleted unused variables to make it easier to read.

   (b) Create a histogram and a table of descriptive statistics for highway fuel efficiency (highway miles per gallon). What is the mean miles per gallon for 2009 four-door vehicles?

   See *ps-3-solutions.xlsx*. From the descriptive statistics, we can see that the mean miles per gallon for 2009 four-door vehicles was 26.3 miles per gallon.

   (c) Create a scatterplot showing the relationship between highway fuel efficiency and car size. Fuel efficiency should be measured on the vertical axis in terms of highway miles per gallon. Car size should be measured by the combined passenger and cargo space.

   See *ps-3-solutions.xlsx*.

   (d) Run a regression of fuel efficiency on car size. Highway miles per gallon should be your dependent variable and total volume (passenger plus cargo space) should be your independent variable. Interpret your regression results. Does the sign of

your coefficient for car size make sense? Would you expect this to be a causal relationship (changes in car size cause changes in fuel efficiency)? What other variables may be influencing the relationship? That is to say, what variables have we not included in our regression that would be correlated with both car size and fuel efficiency?

See *ps-3-solutions.xlsx* for regression results. The coefficient on car volume was -0.128. This suggests that an increase in the size of a car by one cubic foot is associated with a decrease in highway fuel economy of 0.128 miles per gallon. This negative sign makes sense; we expect larger cars to be less fuel efficient. In this particular case, we are not all that interested in the value of the intercept. The intercept would tell us the mileage that a car with zero volume gets which is not a meaningful car size. The R^2 value for the regression was 0.11. This suggests that car volume is explaining only a small portion of the variation in miles per gallon.

There is reason to believe that this could be causal relationship. If we make a car larger it will weigh more and perhaps be less aerodynamic, requiring more fuel power the car.

We could imagine that the engine size and type, the shape of the car, the weight of the car and a variety of other variables would be correlated with both the car size and the miles per gallon to car gets. The coefficient on car volume may be picking up effects of these other variables, not just car size.

(e) Repeat part (d) using engine size rather than car volume as the measure of car size. Engine size is given by the displacement variable and is measured in liters.

See *ps-3-solutions.xlsx* for regression results. From the regression results, we can see that engine size is also negatively correlated with fuel economy as we would expect. An increase in the size of the engine by one liter is associated with a decrease in highway fuel economy of a little over 3 miles per gallon.

(f) Which measure of car size seems to be a better predictor of fuel efficiency? Explain your answer.

The R^2 for the regression that used engine displacement was 0.62. So variation in engine size explains roughly 60% of the variation in highway miles per gallon. This is much better than the regression of highway miles per gallon on car volume which had an R^2 of 0.11. It makes sense that engine size turns out to be a better predictor of fuel efficiency than car volume.

2. **Fuel Economy Over Time** For this question, use the data on fuel economy for both 2009 model cars and for 2001 model cars. These data are contained in the files '2009-

fuel-economy.csv' and '2001-fuel-economy.csv', respectively. The variables are the same in the two files.

(a) We want to look at fuel economy over time. To do this, we'll combine the 2001 and 2009 data into a single dataset. Before we combine the data, we need to construct a new variable in each dataset giving the year. Create a variable named YEAR in each dataset, taking on the value 2001 in the 2001 data and the value 2009 in the 2009 data. Now copy and paste the data from one file into the other so that you have a single dataset with both the 2001 and 2009 data (make certain you line up the variables correctly when pasting the data).

See *ps-3-solutions.xlsx*. Several unused variables have been deleted to make the Excel file easier to read.

(b) For this entire question, we want to focus on a handful of car types and manufacturers and see how fuel economy has changed over time. Sort the data by class and keep the observations for compact, midsize and large cars, deleting all other observations. Now sort the data by manufacturer and keep only the observations for Ford, Chevrolet, Honda, Toyota, BMW and Mercedes. This will leave us with a much more manageable size for the dataset.

See *ps-3-solutions.xlsx*.

(c) First we want to see if fuel economy has improved between 2001 and 2009. Create a table showing the average miles per gallon for city driving in 2001 and in 2009, the average miles per gallon for highway driving in 2001 and in 2009 and the average miles per gallon for combined city and highway driving in 2001 and in 2009 for each car type. Your table should have 18 different miles per gallon numbers in it (3 mpg measures x 3 car types x 2 years). Based on this table, does it look like fuel economy has improved over the past decade?

See *ps-3-solutions.xlsx* for table. For every car type and fuel economy measure, fuel economy has gone down between 2001 and 2009.

(d) Create a scatterplot showing combined city and highway miles per gallon (the COMB MPG (GUIDE) variable) on the vertical axis and engine size, measured by displacement in liters on the horizontal axis. Include two different sets of points on your graph, one set for the year 2001 observations and one set for the year 2009 observations. Based on the scatterplot, does it appear that fuel economy has improved over the past decade?

See *ps-3-solutions.xlsx* for the graph. From the graph, there doesn't appear to be much difference between the relationship between engine size and fuel economy for 2001 and 2009. The drop off in fuel efficiency associated with an additional liter of engine displacement seems to be pretty similar between the two years. So it doesn't seem that the engine technology has gotten less fuel efficient. An engine of a given size appears just as efficient in 2001 and in 2009 from the graph.

(e) Run a regression of fuel economy on engine size, using combined city and highway miles per gallon as your dependent variable and displacement as your independent variable using the year 2001 observations. Run the same regression again but use the year 2009 observations. Based on your regression results, explain whether there is evidence of improvements or declines in fuel economy over the past decade. Consider both the coefficient on displacement and the value for the intercept when coming up with your answer.

See *ps-3-solutions.xlsx* for the regression results. The coefficient on engine displacement comes out much smaller in magnitude for 2009 than for 2001. This means that an increase in engine size is associated with a smaller drop off in fuel efficiency in 2009 than in 2001. In the respect, it appears that engines are getting more fuel efficient over time. The intercept is a bit harder to interpret. If the two regressions had the same slope coefficient, a difference in intercepts would suggest a difference in the average fuel efficiency from one year to the next. However, the smaller intercept for 2009 is a little more difficult to interpret given that the slope coefficients are very different between the two years. What the intercept does tell you is that at an engine size of 0 liters (not a very realistic engine size), a 2009 engine would get fewer miles per gallon than a 2001 engine.

(f) Create a table showing mean displacement by car type for 2001 and 2009. Explain how the values in this table help reconcile differences in your answers for parts (c) and (e).

See *ps-3-solutions.xlsx* for the table. From the table, it is clear that for every class of car engines have gotten substantially larger. This explains why fuel efficieny has gone down even the scatterplot and regression results seemed to suggest that for a given engine size, fuel efficiency has either stayed the same or actually improved since 2001.

(g) The Environmental Protection Agency changed their methods of calculating gas mileage in 2008. According to the EPA, they changed the way they computed gas mileage to take into account the following:

- Faster speeds and acceleration
- Air conditioner use
- Colder outside temperatures

How does this information change your evaluation of how fuel economy has changed between 2001 and 2009?

These changes in the way fuel economy is measured will have major impact on the reported numbers, making the recent numbers lower than the 2001 numbers even for a car for which the 2009 model is identical to the 2001 model. Suppose that the car was measured as getting 30 mpg in 2001. If you take into account air conditioner use, faster speeds, more

acceleration (all things that reduce fuel economy) actually milage will be less than 30 mpg. So the 2009 number, which accounts for these things, will be lower than the 2001 number. The lower fuel economy we observe in 2009 may just be the result of this change in how fueld economy is measured. In fact, if the new way of measuring led to a big drop in fuel economy, it is possible that cars have actually gotten more fueld efficient from 2001 to 2009 even though the reported fuel economy numbers have gone down. To determine what has happened to actual fuel economy over time, we would either convert the 2001 numbers to the new scale or convert the 2009 numbers back to the old scale.