Manning Worthley

ECON 6217

Final Test

Spring 2018

1. <Blank> Time Consuming
2. 1. A dataset of wait-staff employment rates per state and a list of states where the ban has been enacted.
   2. A panel regression could be used to estimate the effect of the tip ban credit. The below model estimates the wait-staff unemployment rate by state *i* and across the time *t* as a function of whether or not that state experienced a ban ( Bani ), an indicator variable for the time the state enacted the ban ( Postt ), and the interaction of the two indicator variables. Also, fixed effects for the state should be added as some states will have different sizes

If b3 is statistically significant then there is an effect from banning the Tip credit.

* 1. One major assumption for the Diff-in-Diff calculation is the assumption for parallel-trends which assumes that the difference in the pre-post treatment is only due to the treatment effect and not due to other sources of variability. In the context of this problem, the assumption is that all else is equal other than the change in treatment.
  2. If, for example, after the tip credit ban is implemented a country wide food shortage occurs, then food prices increase and costs to sell food increases. This in turn, may lead to wait-staff headcount reductions across all states. Estimating the effect of the ban under these circumstances will overestimate the effect of the ban because wait-staff unemployment has increased unrelated to the actual ban’s impact.

1. <Blank>
2. 1. Regression discontinuity models have a key assumption that the pre and post treatment distribution be continuous and linear. The excerpt states that the distribution is not continuous for the migration point system.
   2. Sharp regression discontinuity designs use the jump at the cutoff point as the evidence for discontinuity. \*\*\*\*\*\*\*\*\*Not finished\*\*\*\*\*\*\*\*\*
   3. By using the forcing variable *AGE*, one can see how other forms of death are influenced by age and how the magnitude of the jumps between pre and post. Observing the graph, there is a large jump in MVA deaths when someone turns 21. I am unfamiliar with this dataset as I don’t know if these are directly linked to drinking while impaired or simply due to accident, but I would say that at 21 drinking could cause more vehicular deaths. If the increase in MVA deaths is real and mostly attributed to drinking while impaired, then decreasing the legal drinking age may exasperate the problem of MVA deaths for 19 year old persons.
   4. \*\*\* Not finished \*\*\*\*
3. 1. The propensity score matching model will be able to match a school like Harvard to a comparable for-profit school
4. 1. Like a regression that models the conditional mean, quantile regression models the conditional quantile except using a method that minimizes the absolute deviations from the median. The usefulness of quantile regression on a policy is that the impact of a policy can be measured for different sections of the dataset. For example, one could estimate how the median household income would be affected compared to the lower 5% quantile. This would be important on measuring the impacts only to a specific group of people like the middle class or the very poor. A normal regression would not give the granularity granted by the quantile regression.
   2. As age increases by 1 the median household earnings decreases by 59.85 dollars. As the number of children increases median household earnings decreases by 579.39 for each additional child.
   3. There is a clear pattern for the distribution of age as age increases through the quantiles with a small confidence interval compared to the fixed OLS estimates that cover up the upward trend of age. This suggests that the quantile regression can capture more information than the OLS estimates as the confidence bands for the OLS do not include the trend in age. Religion through the quantiles appears to have an upward trend as well. However, the confidence intervals for both the OLS and the estimated quantiles model are large. In most cases the quantile estimation is in the OLS confidence bands and may not be statistical different from the quantile estimation.
5. <Blank>

1. 1. 1. From the chart of number of tweets over the number of characters, the smart people at Twitter realized certain languages are not able to fully convey a message because users are cutoff by the character limit as there is a high number of tweets capped at 140 characters. To remedy this, Twitter increases the limit of characters to see if the ordered distribution of the characters used in tweets flattens out which would increase the mean of the distribution and signify that English speakers prefer to use more text to convey a message.
      2. A simple causal analysis would be to conduct a difference in the means test on the mean number of characters in tweets in English before and after the implementation of the character limit change. If there is a significant difference, then the English speakers were oppressed by the character limit.
      3. If the objective is to compare English speaking countries..
   2. Correlation does not mean causation. Regression analyses such as simple OLS or panel regression seek to derive whether a variable influences another variable through correlation by scaling a series to minimize the squared errors. Causation analyses can rely on regression analysis to obtain an estimation but implore more theory and tests. For example, the Granger causality test assumes that the past can help predict the future – that the lagged values of variable A can be statistically significant in predicting variable B. If the lagged coefficients of variable A are statistical significant then it is said the variable A Granger causes B.