**ECONOMICS 717**
**APPLIED ECONOMETRICS**
Department of Economics
University of Wisconsin-Madison
Spring 2022
Problem Set: Binary Dependent Variable Models
Version of March 12, 2022
Due on Canvas at 11:00 AM on Monday, March 28, 2022

## Introduction

This problem set uses (a subset of the) data from the Miron and Tetelbaum (2009) *Economic Inquiry* paper that revisits the question of the effect of minimum legal drinking age laws on traffic fatalities. Among other things, Miron and Tetelbaum apply standard panel data methods to U.S. states to examine the effect of variation in the minimum legal drinking age. The paper is available on Canvas if you want to learn more about the underlying analysis. Section II on the historical background of minimum legal drinking ages in the U.S. is particularly relevant.

## Answers

Your write-ups for the problem sets should consist of two portions. The first portion is just the answers to the questions, with whatever text is required to explain them. The second portion, on separate pages, consists of a Stata log file that shows how you got the answers to the empirical questions. The log file must be clear and must include comments that will allow the grader to quickly see the command or commands leading to each answer. It should not include everything you tried – just the final set of commands employed to get the answers. See the syllabus for more details on the format.

## Data set: analysis file extracted from Miron and Tetelbaum (2009)

The data set is called "Economics 717 Miron and Tetelbaum Data" and can be found on the Canvas page. I thank Jeff Miron for supplying the data (with remarkable speed).

It contains the following variables:

state: state number
statename: two-digit postal abbreviation for state name
year: the year
rate18_20ht: traffic fatality rate per 100,000 in the 18-20 age group
mlda: Minimum Legal Drinking Age (MLDA)
mldayr: year in which the state changed its MLDA

In order to focus on a binary treatment, this extract includes only states whose drinking age is always either 18 or 21 (and not ever 19 or 20) over the period from 1975 to 2005, inclusive. It also excludes one state that went from 21 to 18 and then back to 21 over this period.

**Problems**

1. Tell Stata that this is panel data using the command `xtset state year`. [This actually only matters for the bonus problem.]

2. Create a binary treatment indicator called "mlda21" that equals one for state-years in which the MLDA equals 21 and zero for all other state-years.

3. Ignore the panel nature of the data and obtain a naïve estimate of the treatment effect by simply regressing the traffic fatality rate on the treatment indicator. Interpret the resulting estimate and explain why one might not take it seriously.

Be sure to add "`, robust`" to your regression command for this problem and for all of the others too.

[Hint: in responding to questions about taking particular estimates seriously or not, think about where the variation in treatment status comes from and about potential confounders.]

4. Obtain two other naïve treatment effect estimates by repeating the exercise in the preceding problem, first adding only state fixed effects and second adding only year fixed effects. Interpret the resulting estimates and explain why one might not take them too seriously either.

[Hint: if you are new to Stata, it may benefit you to know that, e.g. it translates "`i.year`" in the variable list for a regression translates into a full set of year fixed effects.

5. Obtain the standard estimate of the treatment effect by regressing the traffic fatality rate on the treatment indicator and including both state and year fixed effects. Also, cluster the standard errors at the state level by using the "`cluster(state)`" option along with "`robust`". Interpret your findings. Explain why you might or might not want to take these estimates seriously.

6. Repeat the exercise in the preceding problem but omitting the "`cluster(state)`" option. What happens to the standard errors? Explain.

7. Repeat the exercise in Problem 5 but omitting years of data after 1990. Interpret your findings.

8. Perform a "pre-program test" by coding up a placebo treatment indicator called "placebo82" that equals one in states treated in 1987 (i.e. with "mldayr" equal to 1987) and in years 1982 and later. Then estimate the model in Problem 5 using only the states that always have an MLDA of 21 or that switch in 1987, and using only the years of data prior to 1987. Interpret your findings.

The Miron and Tetelbaum (2009) paper emphasizes that some states "voluntarily" adopted an MLDA of 21 prior to the change in the federal law that "forced" the remainder to do so. Here consider two of those early movers, Michigan (MI) in 1979 and Maryland (MD) in 1983.

9. Separately for MI and MD, estimate the treatment effect of switching to an MLDA of 21 using data for that state and for the "always treated" states. Interpret the resulting estimates. Do these states appear to have distinctive treatment effects?

10. Repeat the exercise in the preceding problem but estimating treatment effects in relative time by replacing the "mlda21" treatment indicator with indicators for the first four years of treatment and remaining years of treatment. Call these variables "mdla21_14" and "mdla_later". Interpret your findings.

Bonus question 1 (not required for full credit). Muck around with the `bacondecomp` user-written Stata command, which implements the decomposition in Goodman-Bacon (2021). You may need to issue the command "`ssc install bacondecomp`" to obtain it. It includes a "help" file. Discuss what you find. In particular, discuss which types of 2 x 2 comparisons these data contain.

Bonus question 2 (not required for full credit). Think about the policy implications of your findings.