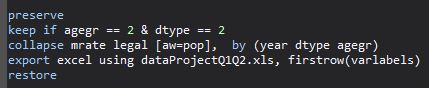
Justin Kacherian

1. ***Create a plot with motor vehicle mortality rate for 18-20 year old’s in the US on the vertical axis and year on the horizontal axis.***

(I collapsed mrate and legal in Stata by year, dtype and agegr, then exported this data to excel to create the plots for questions 1 and 2. The data is weighted by population)



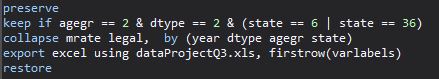
Post 1975, we see a slight increase in average vehicle mortality, which makes sense as many states lowered their MLDA during this period. However, this trend starts to reverse after 1980 and by 1995 average vehicle mortality rates have dropped from an average of roughly 55 deaths per 100,000 to about 35 deaths per 100,000. This is around a 40% decrease in average mortality rate, which is quite substantial. Intuitively, this trend line makes sense as many states opted for a lower MLDA in the 1970’s, however by 1988 all 50 states opted for an MLDA of 21, which suggests the decrease in average mortality rates during this latter time period. States suddenly raising the MLDA to 21 was a direct result of pressure from the government who wanted to punish youthful intemperance by withholding federal highway aid for states with an MLDA of 18.

***2. Overlay the percent of 18-20 year old’s who can legally drink in the US in the first plot.***

When the percent of 18-20 year old’s who can drink is overlaid with the average mortality rate, the interpretation from the previous plot has better context. We can see that one causal interpretation as to why there is a decrease in mortality rates from about 1980 onward, is due to the fact that the average percentage of 18-20 year old’s who could legally drink declined during this period. This makes sense as most states started to increase their MLDA during this time frame.

***3. Same plot as question 1, but with New York and California only.***

(I collapsed the mrate and legal data in Stata by year, dtype, agegr and state, then exported this data to excel to create the plot for question 3)



This plot seems like a classic case where the parallel trends assumption holds and a difference in difference experiment could be done. This is especially intriguing considering California has kept its MLDA at 21 during this entire period and New York lowered its MLDA in 1970 and had not raised it until 1985. A difference in difference experiment could be done where we test whether the treatment effect of lowering MLDA has an effect on average mortality rate.

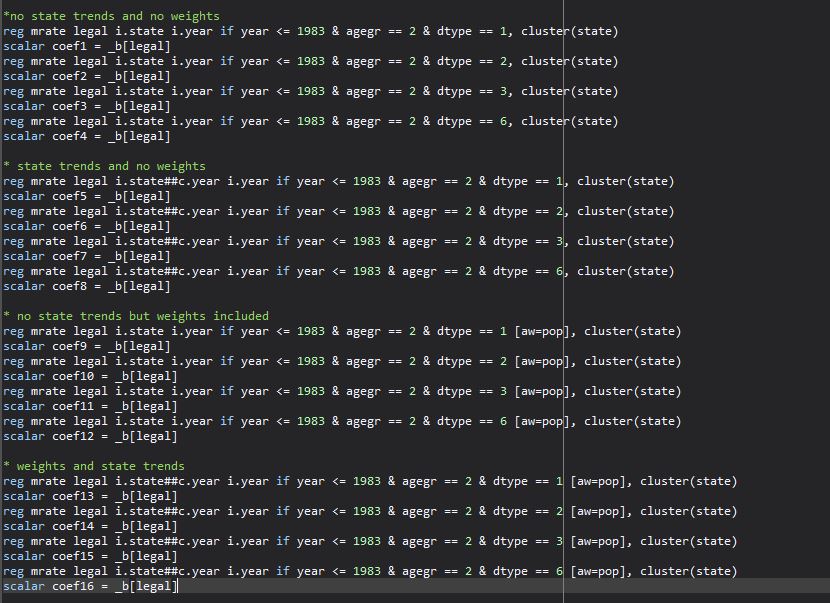
**4. Regress MVM on Legal from 1970-1981 for New York and California only. Include state and year fixed effects. Comment on the coefficient on legal? Why is there no coefficient for state?**

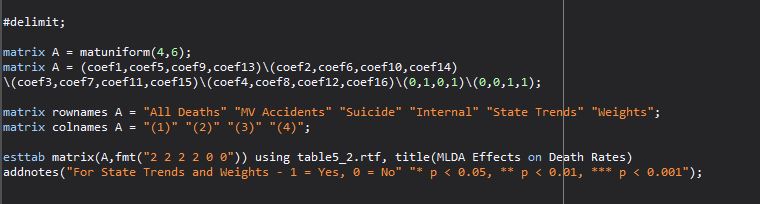
****

Between 1970 and 1981, we see an increase of 7.20 deaths per 100,000 for people in California and New York, which is statistically significant. There is no coefficient for the state of California to avoid multicollinearity.

***5). Replicate the results in Table 5.2 in Mastering Metrics for MVM and all internal causes.***

This table is the result of 16 separate regressions, which depends on the type of death and whether state trends and/or weights are included in the model. I was trying to figure out ways to output the results to a clean formatted table, but I couldn’t figure a way to simply output the 16 stored regressions in the same format they have in the textbook. Instead, I experimented with a matrix. After every regression, I stored the coefficient on legal in a scalar and then created a matrix with all 16 scalars. The resulting matrix was outputted to a word document using the esttab command.





MLDA Effects on Death Rates

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  | (1) | (2) | (3) | (4) |
| All Deaths | 10.80\* | 8.47 | 12.41\*\* | 9.65\* |
| MV Accidents | 7.59\* | 6.64\* | 7.50\*\*\* | 6.46\*\* |
| Suicide | 0.59 | 0.47 | 1.49 | 1.26 |
| Internal | 1.33 | 0.08 | 1.89 | 1.28 |
| State Trends | 0 | 1 | 0 | 1 |
| Weights | 0 | 0 | 1 | 1 |

For State Trends and Weights - 1 = Yes, 0 = No

\* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

***6. Describe the rationale for each of the 4 specifications for MVM?***

Each of the four instances for motor vehicle mortality represent a different regression based on whether state trends and/or weight by population was factored into the model. The first column includes no state trends or weights. The second column includes state trends or more specifically a state-specific linear time trend, which we include by creating an interaction that controls for state and year effects. In the absence of the treatment variable (MLDA), the mortality rates in state “*k”* would follow the linear trend that is captured by the coefficient on this interaction. Including something like this in the model could help uncover the causal effects of an MLDA more clearly as there might be states with mortality rates that were on different trajectories beforehand. So even though the data might show that a drop in the MLDA increases death rates, that particular state might have had a specific linear time trend where their death rates were already increasing regardless. After taking this into account A researcher can then make the proper corrections/assumptions for the magnitude of the causal effect of the MLDA.

Likewise, column 3 weights the model based on the state population. Intuitively this makes sense as a state like California is simply larger than say, Rhode Island and there are presumably more people that drive in the larger state. Likewise, in a small state like Rhode Island with presumably fewer drivers, the data on death rates might be more variable year after year. This further supports the idea of giving a state like this less weight to potentially get more precise estimates.

***7. Comment on the coefficients for MVM and their robustness across models.***

Between 1970 and 1983, we see an increase of roughly 7 deaths per 100,000 in motor vehicle accidents involving people between the ages 18-20. Across all four models, the increase is statistically significant at an alpha of .05. Likewise, when we account for state trends and/or weights based on population, our model does not differ much. As an example, in regression 2 after we account for state-specific trends, we see a decrease in about 1 death per 100,000 for motor vehicle deaths when comparing this to regression 1. Also, we lose some precision in regression 2 as the standard error is slightly higher, however, the causal interpretation in regression 2 supports that of regression 1. Likewise, in regression 3, after accounting for population weights, the model is almost identical to that of regression 1, which does not consider any population based weights. The interpretation across all four models is similar, which is that there is a clear increase in motor vehicle mortality rates among 18-20 year old’s between the years 1970 and 1983.

***8. Why include results for “internal causes?”***

Alcohol is a leading cause of death in the United States. Deaths from internal causes could relate to cancer, disease or even alcohol related causes, which could be influential in the context of this experiment. If the coefficients for internal causes in our model were significant, some of these deaths could be attributed to alcohol related causes. If this were the case there could be bias introduced to the experiment because if people are already dying to alcohol related issues, it would be a fair assumption to state that a higher MLDA would have no effect on these types of people who are presumably alcoholics. It’s important to include these internal causes as an extra measure to ensure that there are no lurking variables that mislead the trend in the data.