

Assignment 3

Ph.D. Applied Microeconometrics
KDI School Fall 2024

2024-10-28

Due date: Monday, November 4th at 6:59pm

For this assignment, you will be using the `castle.dta` data – which I’ve put in the assignment folder – the same data we have been using this week in class.

Below is a list of tasks. I would like you to create a properly formatted PDF file, as if it were a paper. This means that raw code should not appear in the PDF file. When you estimate a regression, the choice of standard errors is up to you. However, I’d like you to justify your choice of standard errors.

Tasks:

1. Create a table with summary statistics *by year*. It should have five columns:
 - The year
 - The proportion of states in the year who have adopted a castle doctrine law (cdl variable)
 - The mean homicide rate (per 100,000 people) for the year
 - The mean population (log) for the year
 - The mean unemployment rate for the year
2. Create a figure that shows adoption of castle doctrine laws over time. I’ll leave the details up to you. Make sure to describe the figure and interpret it.
3. Create a table with four separate regressions:
 - The first regression should be a simple regression of the homicide rate on the castle doctrine law variable.
 - The second regression should be a simple regression of the homicide rate on the castle doctrine law variable and some controls: population (log), income (log), and unemployment rate. Keep in mind that this is at the state level.
 - The third regression should add state fixed effects.
 - The fourth regression should have both state and year fixed effects. This is the “two-way fixed effects” specification we’ve discussed in class.
4. Create a *new* table that includes the same four regressions, but I want you to include some different statistics below the coefficient:
 - p-value with standard errors clustered at the state level
 - p-value calculated using the wild bootstrap, clustered at the state level
5. Discuss the difference in identification strategy across the different columns. How does adding controls and fixed effects change the interpretation of the coefficient on the castle doctrine law variable? What are the assumptions required to interpret the coefficient on the castle doctrine law variable as a causal effect? You do not need to do this for both 3. and 4. Just doing it after 4. is sufficient.
6. Finally, I want you to estimate a causal estimate that is robust to treatment heterogeneity across time. Which you choose is up to you. As a reminder, we used `did2s` in class. How different are the results? Which do you trust the most? Why?

Make sure to *interpret* your output. Do not just present regressions/figures/summary statistics; I also want you to think about what coefficients mean. You can keep your focus on the `cdl` variable, not the other controls in the regression.

As before, please submit the following files:

- Your R Markdown file
- Your knitted PDF file
- Any other scripts you used to complete the assignment