FIN 550: Big Data Analytics - A+ Team

Problem Set #2

V Group Submission. Group member names: Chu-Chun Ku, Hsiang Lee, Yi-Ting Lee

1. DIFFERENCE-IN-DIFFERENCES (30 POINTS)

The data table provides the number of doctors per 100,000 patients in Texas and its neighboring states before and after the implementation of Texas House Bill 4 in 2003. It highlights changes in physician availability over time across these regions.

State	Year	Doctors
Texas	1998	152
Texas	2002	158
Texas	2006	175
Neighbors	1998	196
Neighbors	2002	189
Neighbors	2006	180

- 1. Propose an estimate of the impact of Bill 4 on the number of doctors (per 100,000 patients) using only data for Texas.
 - a. Provide a brief description of your method.

We used the Difference-in-Differences (DiD) method for this question: We compared how the number of doctors changed in Texas and neighboring states (control group) before and after the bill's implementation.

- b. Does this estimate suggest Bill 4 increased or decreased the number of practicing physicians? By how much?
 - In Texas: the number of doctors went up by 17 (175 158).
 - In neighboring states: the number of doctors went down by 9 (180 189).

$$DiD = 17 - (-9) = 26$$

- Therefore, Bill 4 increased the number of doctors in Texas by 26 per 100,000 patients.
- c. Discuss the key assumption required for your estimate to be valid (i.e. no bias).
 We assume that Texas and its neighboring states would have followed parallel trends in the number of doctors if the bill hadn't happened. If this assumption holds true, then any difference observed after the bill can likely be attributed to the bill itself.
- d. Discuss a scenario under which this assumption would be violated.
 - Neighboring states may have implemented other policies or experienced events during this time that influenced the number of doctors, unrelated to Bill 4.
 - Changes in things such as economic conditions might have occurred differently in Texas compared to its neighboring states, which could have potentially affected the results.
- 2. Propose an estimate of the impact of Bill 4 on the number of doctors using only data for 2006.
 - a. Provide a brief description of your method.
 - Using only data for 2006, we can estimate the impact of Bill 4 with a simple treatment/control comparison, comparing TX(treatment group) to neighbors (control group). Estimated treatment effect=175-180= -5 (doctors in TX in 2006 minus doctor in Neighbors in 2006)
 - b. Does this estimate suggest Bill 4 increased or decreased the number of practicing physicians? By how much?

- This estimate suggests that Bill 4 decreased the number of practicing physicians, leading to 5 fewer physicians per 100,000 patients.
- c. Discuss the key assumption required for your estimate to be valid (i.e. no bias).

 The key assumption for this treatment/control comparison to be valid is that the number of practicing physicians in neighboring states equals the number of physicians that would have been practicing in TX, had Bill 4 not been passed because the trend is parallel before the treatment.
- d. Discuss a scenario under which this assumption would be violated.

 The trend is not parallel before the treatment so the key assumption would be violated if the number of physicians practicing in TX differs from the number practicing in neighboring states for any reason other than Bill 4. For example, if incomes in TX were lower than neighboring states, the number of physicians in TX may be lower for that reason. In fact, in 2002 (prior to Bill 4 implementation), the number of physicians in TX and neighbors differed significantly, suggesting that a simple difference between treatment/control likely suffers from selection bias.
- 3. Instead, construct a difference-in-differences estimate of the impact of Bill 4 on the number of doctors.
 - a. Does this estimate suggest Bill 4 increased or decreased the number of practicing physicians? By how much?

Yes, Bill 4 resulted in an increase of 26 physicians per 100,000 patients. DiD = 17 - (-9) = 26

- b. Discuss the key assumption required for your estimate to be valid (i.e. no bias).
 The key assumption is that the number of physicians in Texas and neighboring states would have followed parallel trends in the absence of the policy.
- c. Discuss a scenario under which this assumption would be violated.
 If neighboring states implemented policies aimed at improving physician benefits during the same period, it could affect the parallel trends, thereby violating the key assumption mentioned above.
- d. Set up a test to evaluate whether this key assumption appears to be plausible, using available data. Based on the results of this test, does the assumption appear to be valid?

We evaluate whether the changes in the number of physicians from 1998 to 2002 show parallel trends between the two regions. Texas experienced an increase of 6, while neighboring states experienced a decrease of -7. This indicates that the trends before the policy were moving in different directions, suggesting that the parallel trends assumption may not hold. Incorporating data from a longer time period could help further assess the consistency of the trends.