

Midterm - Fall 2024

ECON 5783 — University of Arkansas

1. A researcher wants to test the folk-wisdom that playing classical music for children benefits their future academic achievement. The researcher surveys parents and records their high-school GPA (Y_i) and whether or not they listened to classical music (D_i) as a child.
 - i. In this context, describe in words what the potential outcomes, $Y_i(0)$ and $Y_i(1)$, are.
 - ii. Do you think the difference-in-means estimator would be biased in this context? If so, which direction do you the treatment effect estimate is biased?
 - iii. Why might a table of summary statistics consisting of sample means of observables, broken down by $D_i = 0$ and $D_i = 1$, be useful in this setting?
2. Say you believe the conditional independence assumption $(Y_i(1), Y_i(0)) \perp\!\!\!\perp D_i \mid X_i$ holds.
 - i. In words, describe what the steps you would take to use a nearest-neighbor matching-based estimator of the treatment effect
 - ii. Say a reader had some concern that overlap would hold in the given application. Why might a balance table be useful for this reader?
3. The simplest treatment effect estimator in a setting where you believe the conditional independence assumption is to difference-in-means estimation for individuals with the same values of $X_i = x$ (e.g. separate estimates for male and females). The overall treatment effect is formed as a weighted average of the estimates $\widehat{\text{CATE}}(x)$.
 - i. Describe why the curse of dimensionality makes this infeasible in many settings.

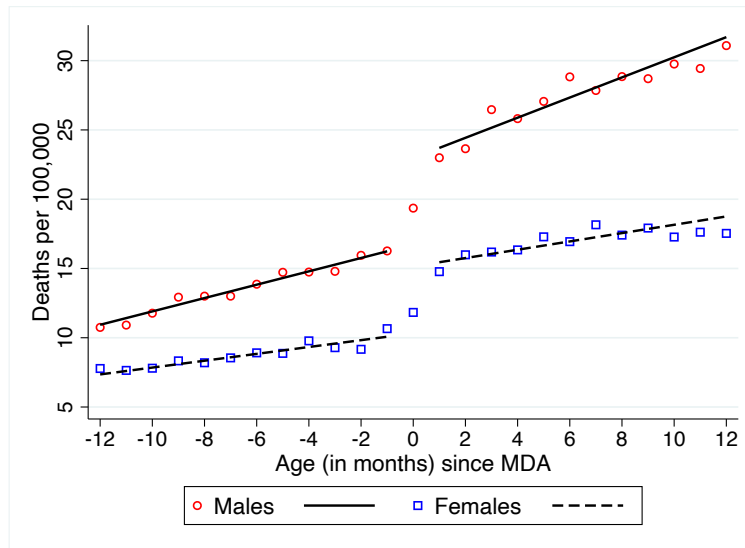
For the following questions refer to Figure 1. These are two figures from the paper “Teenage Driving, Mortality, and Risky Behaviors” in AER: Insights. The authors want to study teenage driving on risky behavior and mortality. They use a state’s minimum driving-age as the cutoff and age as a running variable. Those under the minimum age law can not drive and those above can.

4. Figure 1(a) presents their main results (separately for male and female drivers). They argue that the jump in deaths by motor vehicle accidents is from the driver’s ability to drive. What assumptions do we need to make for this RD to identify the causal effect? I am not looking for

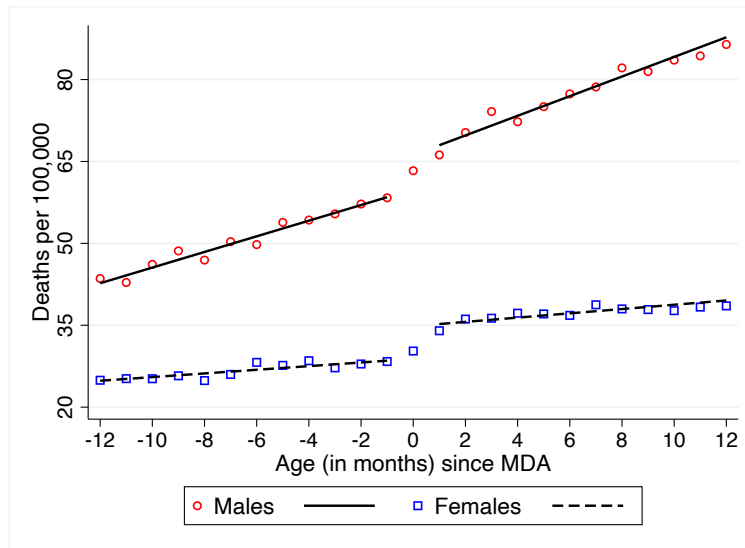
the ‘generic’ assumption here; I want you to describe the assumption in the context of the research question.

5. The authors present figure 1(b) that shows the jump in deaths from ‘Any Cause’ is the same magnitude as the jump *just* from motor vehicle accidents. Do you think this figure helps strengthen their causal claim? Why or why not?

Figure 1 – Results from Huh and Reif (2021, AER Insights)



(a) Deaths by Motor Vehicle Accidents



(b) Deaths by Any Cause