FIN 550: Big Data Analytics – A+ Team

Problem Set #2 – 2 Executive Summary

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

2. TEENAGE DRIVING AND MORTALITY (70 POINTS)

Case Overview

This case looks at the connection between reaching the Minimum Legal Driving Age (MLDA) and teenage mortality rates. It examines whether the differences in mortality rates near the MLDA threshold are linked to driving eligibility and behavior.

Objective

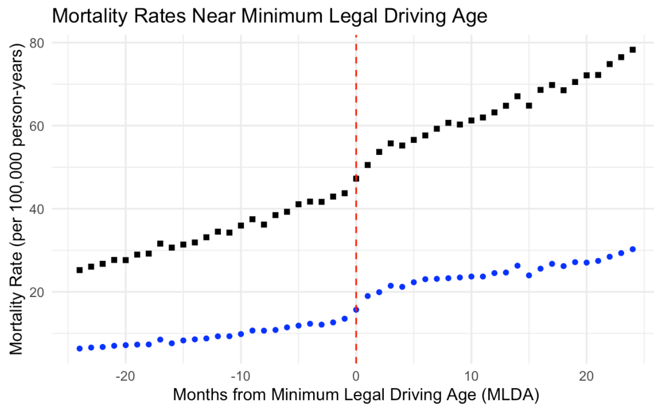
The goal is to understand how driving eligibility affects mortality rates by comparing individuals just above and below the MLDA. It also aims to explore the risks of teenage driving and how driving laws impact public safety.

1. Question 1 : Calculate mortality rates due to any cause for individuals …

   Answer: The descriptive difference shows that the mortality rate is higher for those who reach the MLDA, with a difference of 30.29 per 100,000 person-years. However, this is only observational, does not account for other factors, and therefore cannot confirm causal effects.

   The Regression Discontinuity analysis reveals a significant causal effect of reaching the MLDA on mortality ($p < 0.001$), indicating that the descriptive difference partly reflects the causal impact, though it may overestimate the true effect.

2. Question 2 : Create a scatter plot showing mortality rates…



3. Question 3 : Non-parametric "donut" RD. Calculate a non-parametric RD estimated…

   Answer: As the bandwidth decreases, the estimates for all-cause mortality and motor vehicle accident mortality gradually decrease, while the standard errors tend to increase. The change in precision as the bandwidth becomes smaller is due to two main reasons: (1) Sample size changes: As the bandwidth narrows, the sample size decreases, which causes the estimated treatment effect to become smaller. A larger bandwidth includes more data, allowing the regression model to capture more trends, resulting in a larger estimate. (2) Change in precision: The increase in precision (or standard error) reflects the reduction in sample size, which leads to greater uncertainty in the regression estimates.

Bandwidth <dbl>	RD_All_Cause <dbl>	RD_MVA <dbl>
48	48.83893	21.445071
24	30.27784	15.290869
12	19.06807	11.178446
6	13.16966	8.835371

4. Question 4 : Parametric "donut" RD. Calculate a parametric RD estimated effect…

   Answer: As the bandwidth narrows, point estimates for all-cause and MVA mortality decrease, reflecting reduced bias from distant observations. For example, all-cause mortality drops from 11.89 at 48 months to 6.01 at 6 months. However, smaller bandwidths reduce sample size, increasing standard errors and lowering precision, highlighting the trade-off between bias reduction and estimate reliability. In contrast, parametric RD estimates are significantly lower than non-parametric RD estimates. For example, with a 48-month bandwidth, the parametric estimate for all-cause mortality is 11.89, compared to 48.84 for the non-parametric method. Additionally, parametric RD is more stable across bandwidth changes and has lower standard errors, offering higher precision. Especially in cases with limited sample sizes, parametric methods provide more reliable results.

Parametric Donut RD Estimates with Linear Trends

Bandwidth (Months)	RD Estimate (All–Cause Mortality)	RD Estimate (MVA Mortality)
48	11.891433	9.534183
24	6.877585	6.551764
12	6.611472	5.968510
6	6.012612	4.867183

Conclusion

This study clearly demonstrates that reaching the MLDA leads to a significant increase in teenage mortality rates, particularly those related to motor vehicle accidents. Based on the RD analysis, this impact is consistently observed across various bandwidth conditions. Both non-parametric and parametric estimates indicate a strong causal relationship between driving eligibility and teenage mortality risk. The data shows that as the bandwidth narrows, the precision of the estimates decreases (with higher standard errors), but the bias is reduced, providing a more accurate perspective on the impact of driving regulations. This highlights the need to prioritize stricter driving conditions and enhanced education to mitigate these risks.