

POL-GA 1251
Quantitative Political Analysis II
Homework 5

You are going to replicate some of the results in the following paper,

Anderson, Siwan. 2018. "Legal Origins and Female HIV." *American Economic Review*, 108(6):1407-1439.

The replication material are available on the AER web page as well as our course website.

1. The author's estimation strategy is captured in expression (1) in the paper and the surrounding discussion, with some details (e.g., bandwidth choice) discussed in the results section. Discuss the following. (5 points)
 - What is the "running variable" that the author uses for the regression discontinuity and how is this used in the regression specification?
 - What standard error estimator does the author use and what types of dependencies does this estimator address?
 - What are the bandwidths that the author uses to estimate the effect?
2. Replicate the estimates in Tables 1 and 2 and the graphs in Panels A and B in Figure 3. All of the replication code is available to do so in the author's replication materials. (5 points)
3. Now we are going to use a "local linear regression" approach to estimating these RD effects, working with the `rdrobust` package for R or Stata. (10 points)
 - Re-estimate the effects reported in Tables 1 and 2 using `rdrobust`. Explain what you use for the (i) bandwidth selection procedure (the package allows for options), (ii) kernel, (iii) whether you change the covariate specification, and (iv) your standard error estimator (note that the multi-way cluster estimator may not be available).
 - Create graphs, analogous to Panels A and B in Figure 3, but that display your local linear estimation strategy. The graphs should display a local regression fit to the data (you can plot either raw data or bin-specific means as in the original graph).
 - Discuss how these estimates compare to Anderson's estimates.
4. Finally, we are going to use the "local randomization" approach implemented in the `rdlocrand` package in R and Stata. The key distinctions here are that this approach uses covariate balance as a criterion for selecting a bandwidth and then permutation inference instead of sample-theoretic standard errors. (10 points)
 - Re-estimate the effects reported in Tables 1 and 2 using `rdlocrand` and the `covariates` option for bandwidth selection. Explain what you use for the (i) covariates in the bandwidth selection procedure, (ii) the test statistic for the covariate-driven bandwidth selection procedure, (iii) the kernel for estimating the effects, and (iv) whether you do any covariate adjustment when estimating the effects.
 - Discuss how these results compare to the local linear regression approach and the approach that Anderson used in the paper.