Hans Elliott

EC525 Project II – Impact of China's Huai River Policy on Air Pollution

Ebenstein et al. (2017) estimate the impact of China's Huai River Policy on air pollution and life expectancy. The policy implemented free or highly subsidized coal for indoor heating to the north of Huai River, leading to the construction of coal-powered heating infrastructure in the northern cities, but not the southern.[1] The authors cannot rely on a simple comparison of air pollution in northern and southern cities to measure the causal effect of the policy considering that a variety of differences in demographics, climate, or other factors (both observable and unobservable) exist across the two regions. These variables are likely to introduce bias if they are correlated with a city's location and with a city's air pollution. To overcome this issue, the authors implement a spatial regression discontinuity approach, using the Huai River Boundary as the cutoff in their design and the distance of a city to the boundary as the running variable. This allows them to compare cities that are just north of the river (and thus receiving treatment in the form of the policy) to those that are just south of the river (and not receiving treatment). If identification assumptions are satisfied, the authors have confidence that cities just above and below the river are comparable, so that the southern cities can be considered a control group for the northern, treated cities. This allows them to estimate the discontinuity in pollution at the Huai River Boundary, which they interpret as the causal effect of the policy on pollution (and subsequently, life expectancy).

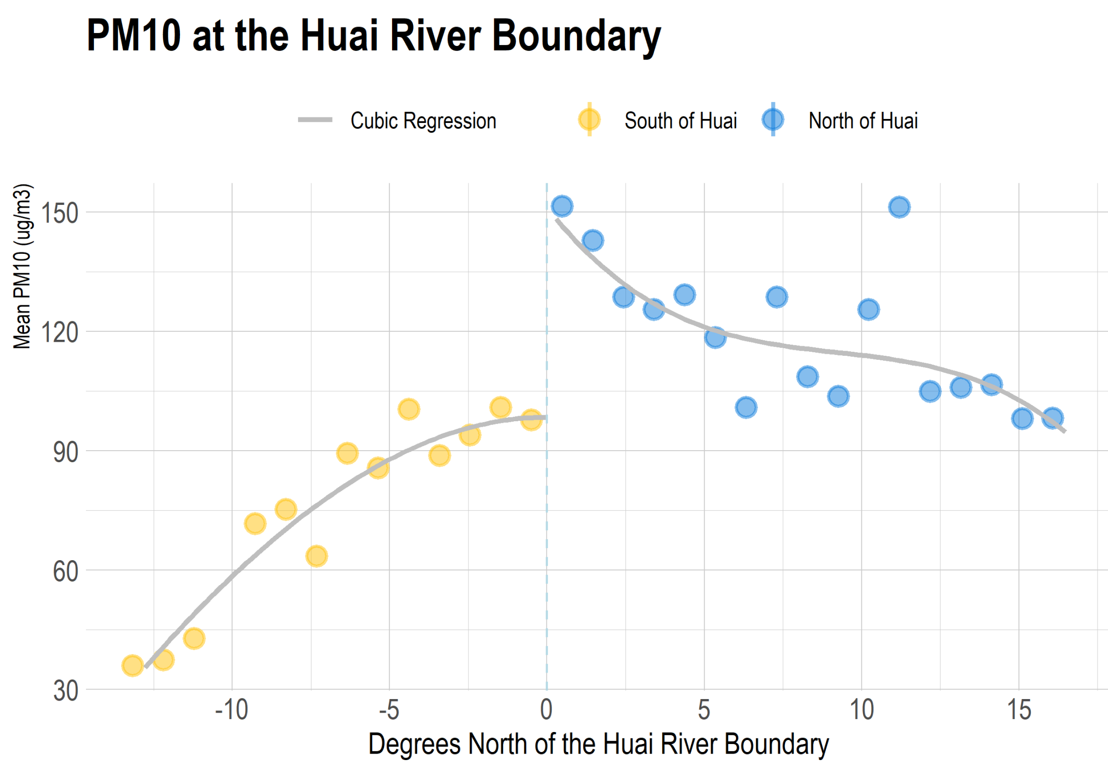
In this report, I replicate the pollution-related results of Ebenstein et al. The outcome variable they use to measure pollution levels (for example, in their Figure 2) is PM10, which is particulate matter smaller than 10 micro-grams (μm).[2] As seen in the replication of Figure 2 below, the outcome variable is

---

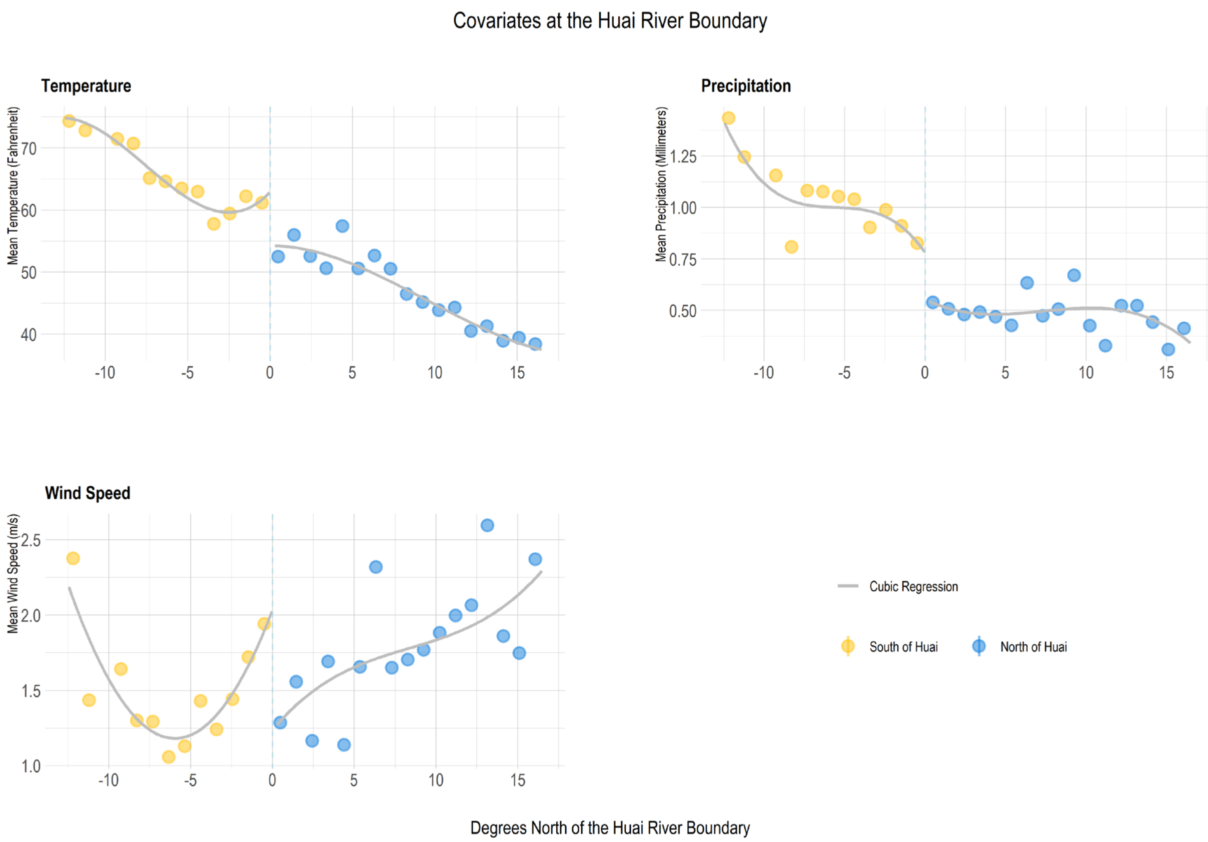
[1] Ebenstein et al., "New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences* 114, no. 39 (2017): 10384-10389.

[2] Ebenstein et al., "New evidence on the impact of sustained exposure to air pollution."

the (average) concentration of PM10 within those cities and the assignment, or running, variable is the cities' (average) distances in latitude-degrees from the Huai River Boundary. This acts as an assignment variable because the Huai River Policy subsidized coal only north of the river, making the boundary a natural discontinuity in treatment (receiving the coal subsidies). The figure below demonstrates that this policy caused a discontinuous change in PM10 concentrations, so that cities to the north of the river (the cities assigned to receive subsidized coal) have significantly higher levels of PM10.



A binned scatterplot, like the one above, collapses the x and y-axis variables into bins, or groups, and plots the mean values for each bin. It is constructed by slicing observations into even-sized bins based on the values of an explanatory variable, summarizing each bin based on the mean value of that explanatory variable, determining the mean value of the outcome variable for each bin, and then plotting the mean outcome variable as a function of the mean explanatory variable. In the following plots, the 161 cities are sliced into 30 bins based on their distance from the Huai River. Each point represents the mean value of the outcome variable plotted by the mean value of the distance variable. This can be interpreted as a non-parametric estimation of the conditional expectation function for the outcome variable given the explanatory variable (distance from the river).



Estimated Discontinuities in Outcomes at the Huai River Boundary

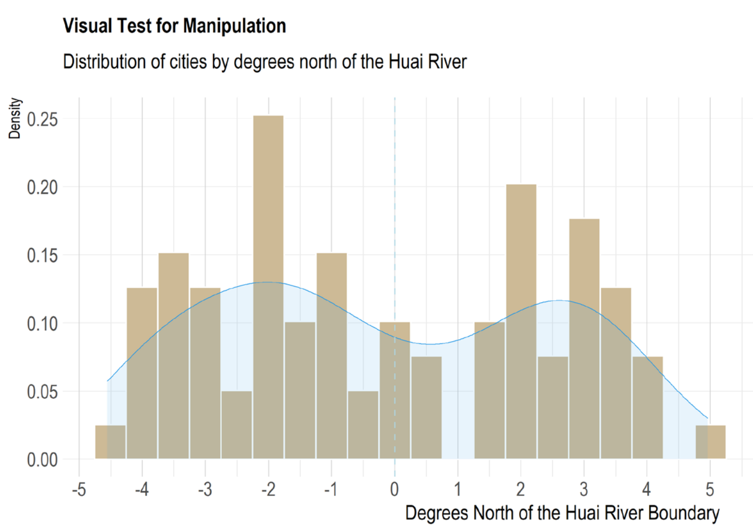
	PM10 (ug/m3)	Temperature (degrees F)	Precipitation (mm)	Wind Speed (m/s)
Point Estimate	67.702	-5.376	-0.140	-0.406
95% Confidence Interval	(16.4, 119)	(-22.27, 11.52)	(-0.36, 0.08)	(-1.31, 0.49)
p-value	0.012	0.535	0.210	0.380
No. Obs. (5 degrees sample)	79	77	77	77

Note: Estimates from cubic regression discontinuity models. Sample limited to 5 degrees north and south of the Huai River. Heteroskedasticity robust SEs used for confidence intervals.

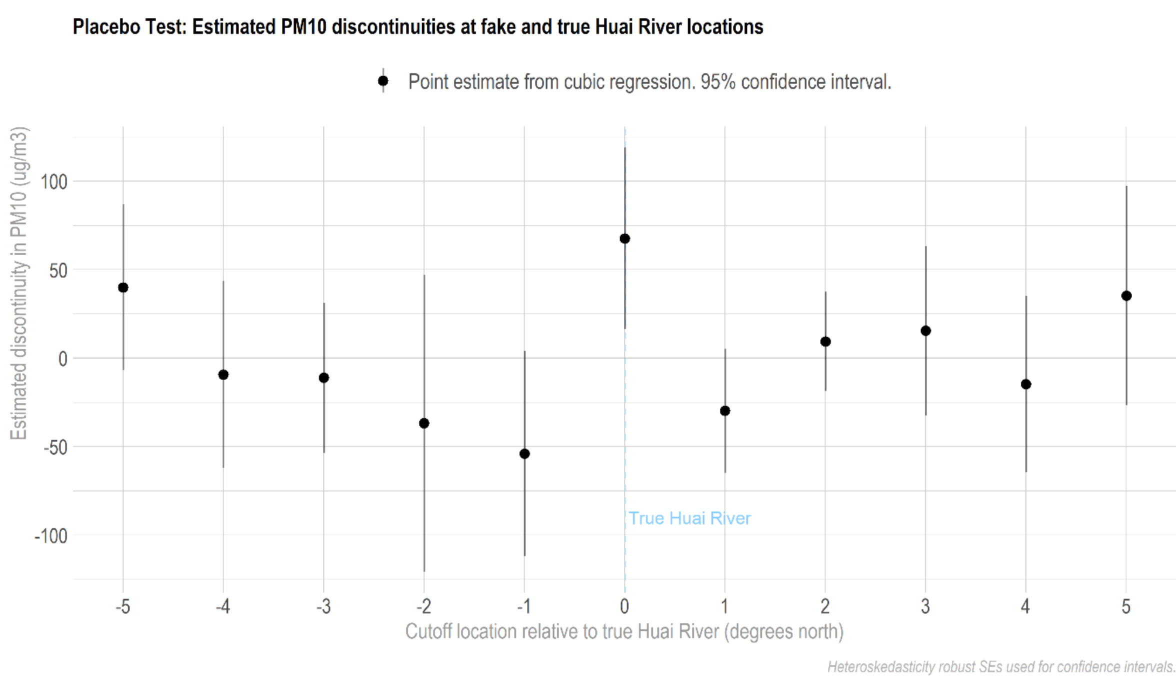
The table above reports the regression results corresponding to the fitted lines in the plots above, estimating the discontinuities in pollution and climate characteristics for cities just north and just south of the Huai River Boundary. The discontinuities are estimated using cubic regression models of the outcome variable on a dummy variable for being north of the Huai River (i.e., a treatment dummy), a variable quantifying the distance of each city from the river in degrees-latitude (i.e., the running variable), one through third degree polynomials of the running variable, and interaction terms between treatment and the polynomials. The models also utilize a bandwidth-like sample restriction to focus the comparisons near the discontinuity. Specifically, only cities within 5 degrees north and south of the Huai River Boundary are included in the model estimation. Heteroskedasticity robust standard errors are used to calculate the confidence intervals for all regressions in this report. The results demonstrate a large, statistically significant increase in PM10 at the Huai River, with a point estimate of 67.7 $\mu g/m^3$ (with a lower bound of 16.4, and an upper bound of 119). Alternatively, there are no statistically significant discontinuities in the other variables around the river.

The graphs and regressions of covariates (temperature, precipitation, and wind speed) attempt to provide insight into the validity of the identification assumption required for a regression discontinuity design to provide a causal estimate. The primary identification assumption in this design is that individuals just above and below the cutoff are statistically identical, so that any discontinuous change in the outcome variable of interest can only be due to the change in treatment that occurs at the cutoff. In other words, to have a valid control group, individuals who are just below the treatment-assigning cutoff must be comparable to individuals just above the cutoff. This means that no other characteristics should change discontinuously at the cutoff. The graphs above demonstrate a covariate smoothness test. By plotting and regressing the covariates as a function of the assignment variable, we can observe any discontinuous changes at the Huai River Boundary (the cutoff). Visually, there do appear to be discontinuities in temperature and precipitation at the river, although they seem to be small in magnitude. This is likely the result of slightly different climates in each region. That said, the regression analysis confirms that there are no statistically significant discontinuities for any of the covariates. Thus, we have some confidence that only PM10 changes discontinuously at the cutoff and that the northern and southern cities are comparable groups.

We can also test the robustness of the identification assumption with a manipulation test. Manipulation occurs when individuals sort themselves at the cutoff of interest so that they may either receive or not receive the treatment. In this case, it is unlikely that we need to worry about sorting considering that the data is at the city-level, and cities cannot relocate easily. It is possible that individuals relocated north to become beneficiaries of the coal subsidization policy or relocated south to avoid the subsequent pollution. However, we are unable to find evidence of this behavior in city-level data. We can confirm that there is no evidence of manipulation by observing in the histogram below that there are no significant discontinuities in the distribution of cities right around the cutoff (i.e., 0 degrees north). There is a slight dip in the distribution around the cutoff, as there are only a few cities located close to the river. However, the distribution is almost symmetric on either side as distance to the river increases, indicating no evidence of sorting.



The authors also test the validity of their design with a placebo test (Figure 4, in their paper). The placebo test attempts to identify any statistically significant discontinuities in PM10 concentrations and life expectancy in other regions of China. By estimating regression discontinuities using fake locations for the Huai River, they can test if the observed difference in pollution concentrations at the real Huai River is a significant finding, or if it is just due to the application of the regression discontinuity design to their specific setting. Their results show that differences in PM10 and life expectancy are only statistically significant at the true Huai River location, suggesting that the differences in PM10 and life expectancy are truly due to the Huai River Policy, and not an artifact of their methodology. The plot below replicates their PM10 placebo tests. The results confirm that the discontinuity in PM10 is only statistically significant from zero at the true Huai River Boundary (0 degrees) since the placebo Huai River estimates include a zero-effect within their 95% confidence intervals. Since the identification assumptions appear to hold, we have some confidence that this design is estimating the true causal effect of the Huai River Policy on PM10.



Reference

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. "New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences* 114, no. 39 (2017): 10384-10389.