Spatial Autocorrelation and Treatment Effect

- Spatial interaction could play a crutial role in policy evaluation
- Spatial interaction may lead to failure of SUTVA assumption
- Spatial autocorrelation may also lead to spurious regressions in which t-stats are inflated
- This can be problematic in a world where economist tend to be obsessed with stars!
- This is the point in Kelly (2019)
- Goal: Examine whether unusually high t statistics combined with severe spatial autocorrelation of residuals in persistence regressions might in fact be connected

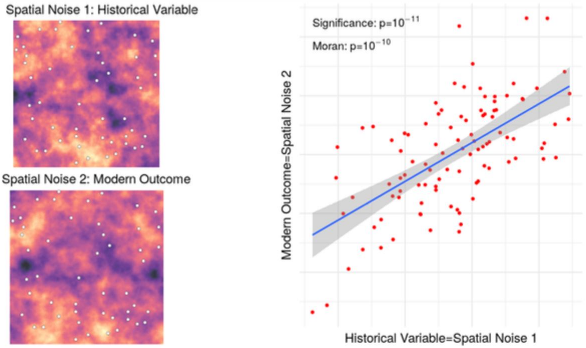
The standard errors of persistence

- Persistence literature: "Many modern outcomes strongly reflect characteristics of the same places in the past"
- This literature usually tries to understand why and how historical factors shape institutions and economic development
- Does the slave trade continue to affect trust between people in Africa? - Nunn and Wantchekon (2011)
- Does a country's prosperity depend on the genetic diversity of its population? - Ashraf and Galor (2013)
- Do arbitrary colonial boundaries continue to drive poverty in Peru and internal conflict in Africa? - Dell (2010) and Papaioannou (2016)
- explanatory power of these persistence variables may easily be overstated because of spatial correlation in error terms

- In a well-behaved regression, residuals should show no pattern,
- In many persistence studies, neighbouring places tend to have similar values of residuals

Spatial noise regressions can appear highly significant

At each town (white dot) we take the value of two noise simulations where dark areas have low values. Call one the modern outcome and the other history, and regress one on the other. The impact of history on the present appears indisputable but is, as the extreme Moran statistic for spatial autocorrelation indicates, spurious.



- Noise simulation: 200 data points in a unit square
- Noise where the correlation of values between points largely disappears after a distance of only 0.1
  - About 20 per cent of synthetic regressions will return a t statistic above 2 (nominal significance level of 5 per cent)
  - correct five per cent significance level is 2.8, which has a nominal p value of 0.005:
  - The significance level is already in error by one order of magnitude.
- If we extend the correlation range to 0.3, nearly half of noise regressions will return a t statistic above 2, and a quarter above 3.
- In order to reach the correct five per cent significance level requires a t of 5.5, something with a nominal significance level of $p = 4 \times 10^{-8}$

Moran Test for Statistical Autocorrelation

- The problem here is that similar units tend to cluster together in space
- Waldo Tobler's (1979) first law: "Everything is related to everything else, but near things are more related to each other"
- Variables that exhibit spatial autocorrelation violate standard assumption of independence across units
- May lead to an under-statement or an overstatement of the importance of the relationship between 2 variables
- One can test for spatial autocorrelation and use estimators that adjust for such autocorrelation
- Moran's I statistic is a good test for presence of autocorrelation
- It is the two-dimensional analogue of the Durbin-Watson test

The Moran's I statistic for spatial autocorrelation is given as:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2} \tag{1}$$

where z_i is the deviation of an attribute for feature i from its mean $(x_i - \bar{X})$, $w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features, and S_0 is the aggregate of all the spatial weights:

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \tag{2}$$

The z_I-score for the statistic is computed as:

$$z_I = \frac{I - \mathrm{E}[I]}{\sqrt{\mathrm{V}[I]}} \tag{3}$$

where:

$$\mathrm{E}[I] = -1/(n-1) \tag{4}$$

$$\mathrm{V}[I] = \mathrm{E}[I^2] - \mathrm{E}[I]^2 \tag{5}$$

- Measures spatial autocorrelation based on both feature locations and feature values simultaneously

How to deal with it?

- Clustering not enough: errors in neighboring clusters will be correlated
- Adjust SE to control for autocorrelation following Conley(1999):
  - Analogous to Adjustment for heteroskedasticity and autocorrelation: consists of calculating a covariance matrix with spatial weights chosen from some kernel (typically uniform kernel: points within a given distance are assigned a weight of unity, and zero otherwise)
  - Most studies set an extremely low cutoff radius so they mistakenly report similar SE than specification without correction
  - Using large cutoffs lead to drop in t-stats but the significance levels they return for finite samples are not entirely reliable and vary according to the spatial clustering pattern of the data

Two-step procedure

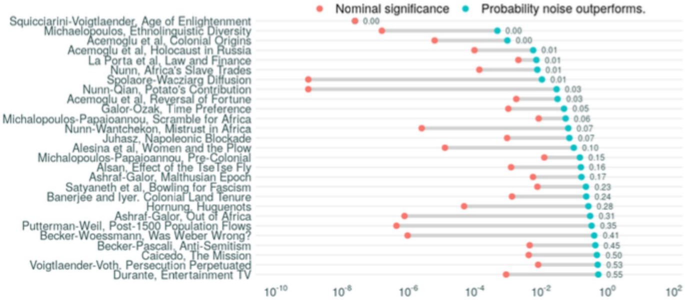
- Compute and report Moran's I statistic:
- Run noise simulation:
  - Replace dependent variable with noise: test the regression's ability to explain what it should not be able to explain) -
  - replace main explanatory variable with noise: By seeing how often random noise outperforms the original persistence variable, we can get some rough idea of its empirical significance level which, as we have seen, can differ by several orders of magnitude from its nominal one.

Moran test for spatial autocorrelation.

This displays the z score of the Moran test for each regression based on the 5 nearest neigbours of each point.

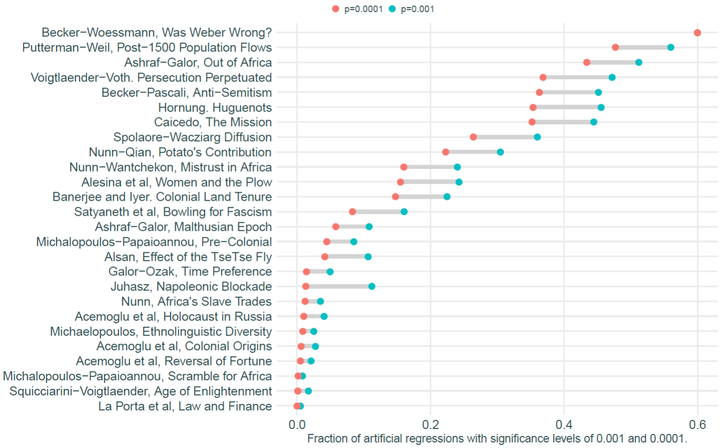
Explanatory power of persistence variables versus spatial noise.

The fraction of artificial regressions where spatial noise has higher explanatory power than the original persistence variable.



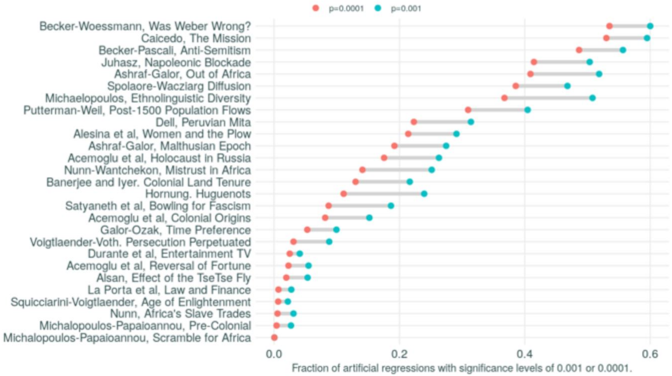
The significance of spatial noise.

The fraction of simulations where a spatial noise explanatory variable was significant at p=0.001 or 0.0001.



Ability of persistence regressions to explain spatial noise.

The fraction of simulations where the persistence variable explained spatial noise at a significance level of 0.001 or 0.0001.



- The effect of spatial auto-correlation in the persistence literature is also a problem for other policy evaluation settings
- DID and RCTs rely on SUTVA assumption
- Dealing with spatial interaction or correlation in treatment or outcome variables is therefore crutial

Treatment Effect Toolbox and Additional Resources

- Scott Cunningham, Causal Inference: The Mixtape: https://mixtape.scunning.com/
- Applied Empirical Methods course by Paul Goldsmith-Pinkham: https://github.com/paulgp/applied-methods-phd

4. Canonical Research Designs
    i. Difference-in-differences
    ii. Event Studies, Synthetic Control + Synthetic DinD
    iii. Instrumental Variables (Part I)
    iv. Instrumental Variables (Part II)
    v. Instrumental Variables (Part III)
    vi. Bartik + Simulated Instruments
    vii. Examiner Designs aka Judge IV
    viii. Regression Discontinuity I: Identification and Groundwork
    ix. Regression Discontinuity II: The Checklist
    x. Regression Discontinuity III: Extensions

Literature on Recent Advances in Applied Micro Methods*

Christine Cai†

September 14, 2021

CLICK HERE FOR THE MOST RECENT VERSION [DARK MODE VERSION]

Contents

*Many of the references listed here are from the applied econometrics courses taught by Michal Kolesár (ECO 539B, Fall 2019, Princeton University) and Arindrajit Dube (ECON 797B, Fall 2020, UMass Amherst). I also added papers that have circulated on #EconTwitter and those that have been suggested to me after I shared the first version of this draft. In particular, I thank Kirill Borusyak, Otavio Canozzi Conceição, Peter Cribbett, Dylan (@dc_pov), Brian Finley, Kevin Grier, Sean Higgins, Jakob Miethe, Tod Mijanovich, Pedro Picchetti, Esteban Quiñones, Jorge Rodríguez, Davud Rostam-Afschar, and Ben Solow for their suggestions. Additional suggestions are welcome ("more is better" ☺).

How to keep up and learn about new methods?

- Engage in discussions with people around you: fellow PhD students, Faculty, etc...
- Go to seminars, workshops, conferences, reading groups: meet with speakers to discuss your projects

Be curious!

Evaluation

- PS with data in pairs of two
- Referee report (individual)
  - Incentives for hierarchy in comments: Major versus minor comments
  - Incentives for submitting major comments that are relevant and not reported by other students
- Weight of each assignment will be announced when you get first assignment