

Exercise A

Assume you want to learn more about the spatial dimension of the growth rate of productivity. Find the relevant files (a the binary RData file `data1.rda`, and a Shapefile `EU27.*`) on Canvas. Open both files in R; focus on the eloquently named variables

- `pr80b` and `pr103b`, which hold a region's productivity in 1980 and 2003,
- `lninv1b`, which holds the (log of) investment, and
- `lndens.empb`, which holds the (log of) density of employment.

Restrict your analysis to Austria, Germany, Italy, France, Spain and Portugal for convenience and exclude overseas territories if necessary.

- Calculate the growth rate of productivity from 1980 to 2013 and create a map that shows the productivity growth for each region.
- Generate three different spatial weights matrices using (i) a distance threshold, (ii) smooth distance-decay, and (iii) a contiguity-based measure.
 - Compare the matrices; use you knowledge of graph theory and linear algebra.
 - Plot the resulting matrices as raster images (i.e., directly).
 - Try and visualize the network that they represent.
- Compute a suitable measure of spatial autocorrelation for productivity growth using these matrices. Point out differences, if there are any.
- Estimate a linear regression model using OLS, where you regress the growth of productivity on the initial productivity in 1980, the investment, and the employment variable. Check for spatially autocorrelated errors (visually and statistically), and compare them across the different weighting schemes. Briefly discuss potential implications of spatial dependence.

Exercise B

Read the paper by [Caicedo \(2019\)](#)¹ on the long-term effects of Jesuit missions on socioeconomic outcomes in South America (focus on Sections I, III, and IVa). Load the provided dataset `literacy_Arg-Bra-Par.dta`.

- Create a map using some of the variables of interest. You can find shapefiles (and other spatial data formats) for sub-national administrative units [here](#). Add borders of the other states in the three countries analyzed and national borders for all countries.
- Reproduce the main results in Table 2 of [Caicedo \(2019\)](#). Try to replicate the results using Conley standard errors.
- How does Valentian Caicedo operationalize distance? Can you think of other ways that might make sense? Plot three potential functions and discuss whether and how they might affect estimation results.
- Explain what role geography and other spatial variables might play in regression that relates variables with a long time gap between them. Think of potential confounders and how spatial dependence may bias estimates.

¹Valencia Caicedo, F. (2019). The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America. *Q. J. Econ.*, 134(1), 507–556.

Exercise C

Recall ‘The perils of peer effects’ (Angrist, 2014)². Write a short text (not more than 800 words) on the ‘The perils of ignoring peer effects’.

- Touch on the topics of drawing valid inference and the trade-off between internal and external validity (think of an experimental settings vs., e.g., an actual classroom), and the goals of (applied and methodological) scientific research.
- Briefly explain how network dependence (spatial, social, etc.) may impact *validity* and *relevance* of a certain instrument. Consider weather instruments, the quarter of birth instrument by Angrist and Krueger (2001)³, or some instrument that you are familiar with as an example.

Exercise D

What is the format of this image?⁴



²Angrist, J. D. (2014). The Perils of Peer Effects. *Labour Economics*, 30, 98–108.

³Angrist, J. D., & Krueger, A. B. (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives* 15 (4): 69–85. <https://doi.org/10.1257/jep.15.4.69>.

⁴It's a raster format, so it's relevant to the course.