

SPATIAL DATA

Valerio Leone Sciabolazza

valerio.leonesciabolazza@uniroma1.it

This course introduces students to the use of spatial data and econometric tools for causal inference. Spatial data have become a powerful resource for economists and policy analysts because they allow us to measure treatment exposure, identify valid sources of exogenous variation, and detect spillovers across geographic units.

Throughout the course, we will explore how spatial data can be employed to answer core policy questions—for example: How does infrastructure shape economic development? What are the long-term effects of historical shocks? How can environmental or geographic features serve as instruments to identify causal effects?

Each lecture combines the study of empirical papers with hands-on exercises in R. Students will replicate key research designs, applying spatial tools to concrete problems of measurement, identification and evaluation. While AI tools can assist with these tasks, understanding the underlying methods and their limitations remains essential for credible empirical analysis. By the end of the course, students will be able to design and implement research strategies that leverage spatial data to evaluate policies and interventions.

Syllabus

Lesson 1 – *Introduction to Causal Inference and Spatial Data* – 03/05/27 16:00-18:00.

In this lecture, we will discuss how satellite image data have recently been used by economists and learn the basics of spatial datasets, including polygons, raster data, and map projections. We will also discuss how spatial data are constructed and used in empirical research, and their main sources of measurement error, which is often non-classical and can bias estimated effects in either direction.

Reference papers: Alix-Garcia, Jennifer, and Daniel L. Millimet. 2023. “Remotely incorrect? Accounting for nonclassical measurement error in satellite data on deforestation.” *Journal of the Association of Environmental and Resource Economists*, 10(5): 1335-1367; Donaldson, Dave, and Adam Storeygard. 2016. “The View from Above: Applications of Satellite Data in Economics.” *Journal of Economic Perspectives*, 30(4): 171–198; Kudamatsu, Masayuki. 2018. “GIS for Credible Identification Strategies in Economics Research.” *CESifo Economic Studies*, 64(2): 327-338; Wuepper, David, Wyclife Agumba Oluoch, and Hadi Hadi. 2025. “Satellite data in agricultural and environmental economics: Theory and practice”, *Agricultural Economics*, 56(3): 493-511.

Module 1: Geography as a Source of Exogenous Variation

Geographic and environmental conditions can provide variation useful for identifying causal effects. We will examine how to incorporate this variation into empirical research designs and assess the assumptions required for a causal interpretation.

Lesson 2 – *Spatial join* – 05/05/27 14:00-16:00.

We will replicate how this paper spatially merges ethnic group level data with the grid of weather data, a great source of exogenous variation. We will also discuss how the resolution and aggregation of weather data affect measurement and the results of the analysis.

Reference paper: Alsan, Marcella. 2015. “The Effect of the TseTse Fly on African Development.” *American Economic Review*, 105(1): 382–410.

Also relevant: Josephson, Anna, Jeffrey D. Michler, Talip Kilic, and Siobhan Murray. 2026. “The mismeasure of weather: Using earth observation data for estimation of socioeconomic outcomes.” *Journal of Development Economics*, 178: 103553.

Lesson 3 – *Distance* – 07/05/27 11:00-13:00.

This lecture demonstrates how Nunn (2008) constructs instrumental variables using distance from each country’s centroid to the nearest coastal point. We will also discuss how the choice of reference point and distance metric shapes the instrument, and why geographic proximity alone does not guarantee that the exclusion restriction holds.

Reference paper: Nunn, Nathan. 2008. “The Long-Term Effects of Africa's Slave Trades.” *Quarterly Journal of Economics*, 123(1): 139–176.

Lesson 4 – *Elevation* – 10/05/27 16:00-18:00.

We will replicate how this paper uses river slope to construct instrumental variables for dam construction, illustrating how terrain characteristics can provide exogenous variation in infrastructure placement.

Reference paper: Duflo, Esther, and Rohini Pande. 2007. “Dams.” *Quarterly Journal of Economics*, 122(2): 601–646.

Module 2: Defining Spatial Units for Causal Identification

The choice of spatial units and the construction of variables shape the comparisons underlying empirical analysis. We will examine how to define these units and combine spatial measures to address questions of measurement and causal identification.

Lesson 5 – *Zonal Statistics* – 12/05/27 14:00-16:00.

We will replicate how this paper calculates the standard deviation of land suitability for agriculture at the 0.5-degree cell level within each “virtual country”, where the boundary is exogenously defined. We will also discuss how the choice of spatial units and aggregation affects measurement and the results of the analysis.

Reference paper: Michalopoulos, Stelios. 2012. “The Origins of Ethnolinguistic Diversity.” *The American Economic Review*, 102(4): 1508–1539.

Lesson 6 – *Map Algebra* – 14/05/27 11:00-13:00.

Map algebra allows us to combine raster data at the cell level to construct measures of agricultural productivity and other geographic characteristics. We will replicate how this technique is used to examine the role of land productivity and crop appropriability in the origins of states.

Reference paper: Mayshar, Joram, Omer Moav, and Luigi Pascali. 2022. “The Origin of the State: Land Productivity or Appropriability?” *Journal of Political Economy*, 130(4): 1091-1144.

Module 3: Modeling Treatment Exposure

A recurring challenge with spatial data is that observations are rarely independent across space, and treatments may generate spillover effects on neighboring units. We will examine methods for modeling spatial autocorrelation both in observational settings and in quasi-experimental designs, leveraging the fact that treatment effects often decay with distance.

Lesson 7 – *Buffer* – 17/05/27 16:00-18:00.

We will replicate how Conley and Udry (2010) match each agricultural plot with its neighboring plots within a radius of one kilometer in order to control for common shocks in peer effect estimation.

Reference paper: Conley, Timothy G., and Christopher R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review*, 100(1): 35–69.

Also relevant: Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of econometrics*, 92(1): 1-45; Manski, Charles F. 1993. "Identification of endogenous social effects: The reflection problem." *The review of economic studies*, 60(3): 531-542.

Lesson 8 – *Spatial Regression Discontinuity Design* – 19/05/27 14:00-16:00.

We will replicate how Dell (2010) obtains each observation's distance to the treated area boundary (i.e., the cutoff) and the nearest boundary segment in order to implement a regression discontinuity design (RDD). In addition, we will replicate how the paper accounts for elevation in calculating road length in Peru, a hilly country. The lecture will also cover recent methodological advances in RDD, focusing on strategies to improve identification around cutoffs and accurately estimate changes in outcomes.

Reference paper: Dell, Melissa. 2010. "The Persistent Effects of Peru's Mining Mita." *Econometrica*, 78(6): 1863–1903.

Also relevant: Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocio Titiunik. 2019. "Regression discontinuity designs using covariates." *Review of Economics and Statistics*, 101(3): 442-451.

Lesson 9 – *Spatial first differences, Ring-based approaches* – 21/05/27 11:00-13:00.

We will examine methods for modeling spatial autocorrelation both in observational settings (e.g., Conley's standard errors) and in quasi-experimental designs (e.g., spatial first differences, ring-based approaches) leveraging the fact that treatment effects often decay with distance.

Reference Papers: Butts, Kyle. 2023. "Difference-in-differences with geocoded microdata." *Journal of Urban Economics*, 133, 103493; Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of econometrics*, 92(1): 1-45; Druckenmiller, Hannah, and Solomon Hsiang. 2018. "Accounting for unobservable heterogeneity in cross section using spatial first differences." NBER working paper series, No. w25177.

EXAM - TBD

The final exam will consist of an applied exercise. Students will be assigned excerpts from a recent empirical paper describing the setting, research question, and data. Based on this information, students will be required to design an identification strategy that leverages the methods learned throughout the course. The exam will also require a critical discussion of the proposed approach, including its strengths and limitations. The exam will be held in person and will have a duration of two hours.

Required materials:

In order to participate to the class, you'll need the following:

- Your own laptop.
- R and RStudio installed on your laptop (please make sure to have the latest available version).
- The relevant materials, provided during the first class, downloaded to a folder on your laptop. Please make sure you have an internet connection.
- The following R packages installed: binsreg, conleyreg, cubelyr, data.table, devtools, dplyr, elevatr, exactextractr, fixest, furr, ggplot2, ggrepel, ggthemes, glue, haven, kfbmisc, KernSmooth, lpridge, lwgeom, modelsummary, np, parallel, patchwork, pbapply, purrr, raster, rdrobust, readr, readxl, reshape2, scales, sf, sp, stargazer, stars, starsExtra, stringr, terra, tigris, tidyr, tidyverse, treat, viridis, WDI.

ADDITIONAL READINGS

Lesson 1 – *Introduction*

Gibson, John, Susan Olivia, Geua Boe-Gibson, and Chao Li. 2021. "Which Night Lights Data Should We Use in Economics, and Where?" *Journal of Development Economics*, 149:102602.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review*, 102(2):994–1028.

Lesson 2 – *Spatial join*

Axbard, Sebastian. 2016. "Income Opportunities and Sea Piracy in Indonesia: Evidence from Satellite Data." *American Economic Journal: Applied Economics*, 8(2):154–194.

Bai, Ying, and Ruixue Jia. 2016. "Elite Recruitment and Political Stability: The Impact of the Abolition of China's Civil Service Exam." *Econometrica*, 84(2):677–733.

Feyrer, James, and Bruce Sacerdote. 2009. "Colonialism and Modern Income: Islands as Natural Experiments." *Review of Economics and Statistics*, 91(2):245–262.

Hoxby, Caroline M. 2000. "Does Competition among Public Schools Benefit Students and Taxpayers?" *American Economic Review*, 90(5):1209–1238.

Lesson 3 – *Distance*

Bekker, Paul A. 1994. "Alternative Approximations to the Distributions of Instrumental Variable Estimators." *Econometrica*, 62(3):657–681.

Gebresilasse, Mesay. 2023. "Rural Roads, Agricultural Extension, and Productivity." *Journal of Development Economics*, 162:103048.

Moreira, Marcelo J. 2003. "A Conditional Likelihood Ratio Test for Structural Models." *Econometrica*, 71(4):1027–1048.

Oster, Emily. 2012. "HIV and Sexual Behavior Change: Why Not Africa?" *Journal of Health Economics*, 31(1):35–49.

Woodruff, Christopher, and René Zenteno. 2007. "Migration Networks and Microenterprises in Mexico." *Journal of Development Economics*, 82(2):509–528.

Lesson 4 – *Elevation*

Chen, Ting, and James Kai-sing Kung. 2016. "Do Land Revenue Windfalls Create a Political Resource Curse? Evidence from China." *Journal of Development Economics*, 123:86–106.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000." *American Economic Review*, 106(3):479–524.

Dinkelman, Taryn. 2011. "The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review*, 101(7):3078–3108.

Gagliarducci, Stefano, Massimiliano Gaetano Onorato, Francesco Sobbrino, and Guido Tabellini. 2020. "War of the Waves: Radio and Resistance during World War II." *American Economic Journal: Applied Economics*, 12(4):1–38.

Qian, Nancy. 2008. "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance." *Quarterly Journal of Economics*, 123(3):1251–1285.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics*, 125(3):1253–1296.

Yanagizawa-Drott, David. 2014. "Propaganda and Conflict: Evidence from the Rwandan Genocide." *Quarterly Journal of Economics*, 129(4):1947–1994.

Lesson 5 – *Zonal Statistics*

Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. 2016. "Ethnic Inequality." *Journal of Political Economy*, 124(2):428–488.

Baskaran, Thushyanthan, Brian Min, and Yogesh Uppal. 2015. "Election Cycles and Electricity Provision: Evidence from a Quasi-Experiment with Indian Special Elections." *Journal of Public Economics*, 126:64–73.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber. 2012. "The Political Economy of Deforestation in the Tropics." *Quarterly Journal of Economics*, 127(4):1707–1754.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century." *American Economic Journal: Macroeconomics*, 4(3):66–95.

Easterly, William, and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics*, 112(4):1203–1250.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review*, 102(2):994–1028.

Hodler, Roland, and Paul A. Raschky. 2014. "Regional Favoritism." *Quarterly Journal of Economics*, 129(2):995–1033.

Martínez, Luis R. 2022. "How Much Should We Trust the Dictator's GDP Growth Estimates?" *Journal of Political Economy*, 130(10):2731–2769.

Michalopoulos, Stelios, and Elias Papaioannou. 2013. "Pre-Colonial Ethnic Institutions and Contemporary African Development." *Econometrica*, 81(1):113–152.

Michalopoulos, Stelios, and Elias Papaioannou. 2014. "National Institutions and Subnational Development in Africa." *Quarterly Journal of Economics*, 129(1):151–213.

Montalvo, José G., and Marta Reynal-Querol. 2005. "Ethnic Polarization, Potential Conflict, and Civil Wars." *American Economic Review*, 95(3):796–816.

Lesson 6 – *Map Algebra*

Asher, Sam, and Paul Novosad. 2020. "Rural Roads and Local Economic Development." *American Economic Review*, 110(3):797–823.

Axbard, Sebastian. 2016. "Income Opportunities and Sea Piracy in Indonesia: Evidence from Satellite Data." *American Economic Journal: Applied Economics*, 8(2):154–194.

Lesson 7 – *Buffer*

Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, 72(1):159–217.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. 2015. "The Housing Market Impacts of Shale Gas Development." *American Economic Review*, 105(12):3633–3659.

Lesson 8 – *Spatial Regression Discontinuity Design*

Berger, Melissa, Gerlinde Fellner-Röhling, Rupert Sausgruber, and Christian Traxler. 2016. "Higher Taxes, More Evasion? Evidence from Border Differentials in TV License Fees." *Journal of Public Economics*, 135:74–86.

Jones, Maria, Florence Kondylis, John Loeser, and Jeremy Magruder. 2022. "Factor Market Failures and the Adoption of Irrigation in Rwanda." *American Economic Review*, 112(7):2316–2352.

Keele, Luke J., and Rocío Titiunik. 2015. "Geographic Boundaries as Regression Discontinuities." *Political Analysis*, 23(1):127–155.

Michalopoulos, Stelios, and Elias Papaioannou. 2014. "National Institutions and Subnational Development in Africa." *Quarterly Journal of Economics*, 129(1):151–213.

Sokoloff, Kenneth L., and Stanley L. Engerman. 2000. "Institutions, Factor Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives*, 14(3):217–232.

Thistlethwaite, Donald L., and Donald T. Campbell. 1960. "Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment." *Journal of Educational Psychology*, 51(6):309–317.

Lesson 9 – *Spatial first differences, Ring-based approaches*

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf. 2015. "The Economics of Density: Evidence from the Berlin Wall." *Econometrica*, 83(6):2127–2189.