Introduction to Economic Geography

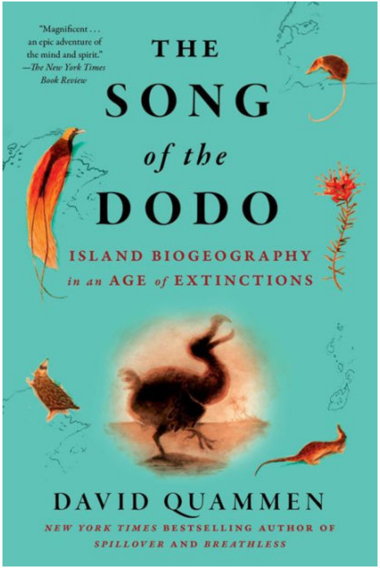
Bruno Barsanetti

Economic Geography
FGV EPGE

Economic geography is highly concentrated

- *H. sapiens* habitat restricted to water-free surfaces (29% of the Earth)
- Population (and economic activity) is highly concentrated within cities
  - 1-3% of available area, but 56% of the population $\rightarrow$ focus of much spatial economic
  - 44% for agriculture (only $\frac{1}{3}$ is cropland) $\rightarrow$ it is also relevant
- Broader research question: **what does explain this distribution?**

A point from biogeography: learning from "islands"



This class

Unit 1: Core-periphery model

- agglomeration economies + transportation costs
- explain concentration, multiple equilibria, symmetry breaking

Unit 2: History versus expectations

- path dependence
- how economic geography changes, why, which shocks
- evidence from historical experiments

Unit 3: Quantitative spatial models

- trade-influenced structural approach
- “realistic geography”

Unit 4: Role of cities in economic geography

- system of cities models
- urban type (function) and size
- brief overview of main urban economics topics

Unit 5: Internal geography of cities

- Alonso-Muth-Mills model $\rightarrow$ monocentric cities
- land use, city shape, housing informality

Space in economics: its origins

- von Thunen's *The Isolated State* (1826)
  - transport costs influence rents and land use
- Central place theory (Losch, Christaller), regional science (Isard)
- Space can be explained by general equilibrium (Debreu, 1959)
  - but mostly uninteresting
  - CRS + fixed factors $\rightarrow$ determinate equilibrium
- Urban economics:
  - monocentric city model of Allonso-Muth-Mills $\rightarrow$ commuting + construction
  - Henderson (AER 1974), "The sizes and types of cities"
  - investigation of agglomeration economies, (re)discovery of Marshall and Jacobs
  - from labor: Rosen-Roback model, regional evolutions (Blanchard and Katz, 1992)

Space in economics: its recent past

- New Economic Geography
  - increasing returns and transportation costs
  - combines tools of urban (Fujita) and trade (Krugman)
  - stylized environments, usually simple geography
  - key predictions: multiple equilibria, path dependence
- Empirical evidence based on natural experiments
  - Davis and Weinstein (AER 2002): historical data, large shocks
  - Bleakley and Lin (QJE, 2012): evolving fundamentals
  - ...
- Quantitative spatial economics (QSE)
  - extends structural trade models to richer geographies
  - sufficient statistics (Donaldson)

Space in economics: its present and future

- Growing research field in economics
- Fast growth of QSE papers
  - expanding computing power
- Amazing spatial data sources
  - satellite data: nightlights, daylight images
  - georeferencing historical maps and administrative data
  - GPS information, cell-phone data, ...

The global distribution of economic activity: Nature, history, and the role of trade

Henderson, Squires, Storeygard, Weil

Quarterly Journal of Economics

2018

Motivation

- First nature: physical characteristics of different sites, geographic fundamental
- Does the importance of first nature change over time?
  - paper focus: agricultural suitability $\times$ access to trade
  - see also Alix-Garcia and Sellars (JUE 2020), “Locational fundamentals, trade, and the changing urban landscape in Mexico”
- But if they do, and there is some path dependence, then the importance of some geographic fundamentals will be different in richer economies (that urbanized and industrialized earlier) than in poorer ones

Changes over time

- In the past 200 years, transportation costs have fallen due to more modern transportation technologies
- Moreover, international trade grew faster than domestic economies $\rightarrow$ access to international markets has become more important
- This paper: show these different patterns
  - note, again, that we have a paper that focuses on some heterogeneity

Data

- Dependent variable: log of nightlights in 2010
  - 2010 Global Radiance Calibrated Nighttime Lights
  - collapsed into a 0.25x0.25-degree grid (770 sq km at Equator)
  - replaced zeros by small values
- Base covariates:
  - Stability of malaria transmission (Kiszweski et al 2004)
  - Terrain ruggedness (affects trade and ag)
- Agriculture covariates:
  - average temperature and precipitation
  - length of growing period (FAO/IIASA)
  - land suitability (Ramankutty et al 2002): prob. of agriculture given soil and climate
  - elevation
  - biome classification
- Trade covariates:
  - distance to nearest coast, navigable river, major lake, natural harbor

Demeaned log nightlights

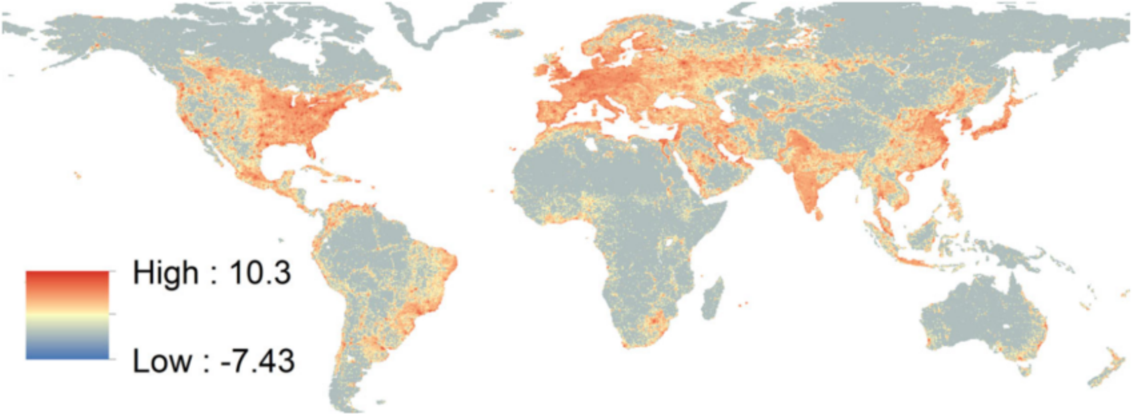
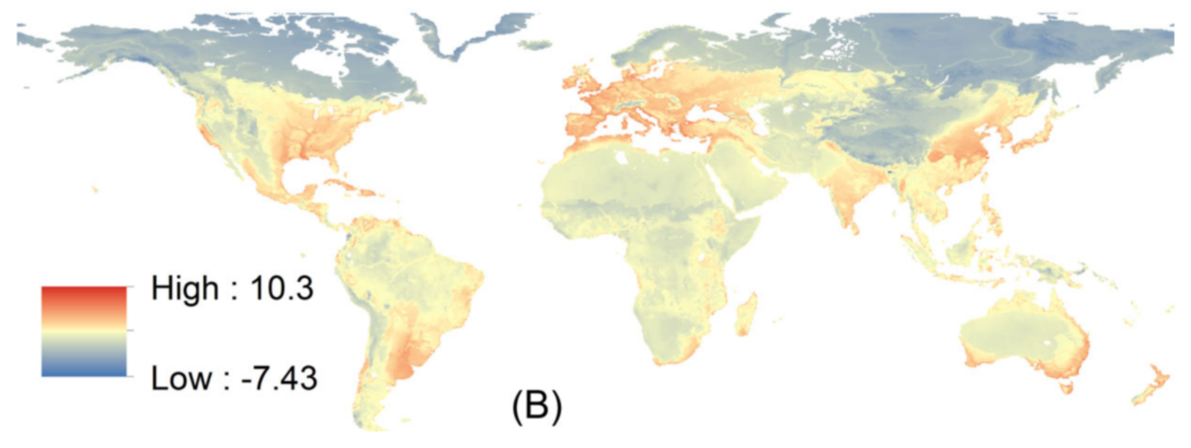


FIGURE I

Demeaned ln(lights)

Predicted nightlights (pooled, FEs in estimation but not in the prediction)



Relative importance: homogeneous equation

	No country FEs (1)	With country FEs (2)
Panel A: R^2		
All variables ($N = 242{,}184$)	0.467	0.577
Base variables (malaria, ruggedness)	0.020	0.355
Agriculture variables (plus base)	0.450	0.566
Trade variables (plus base)	0.066	0.370
Country fixed effects		0.345

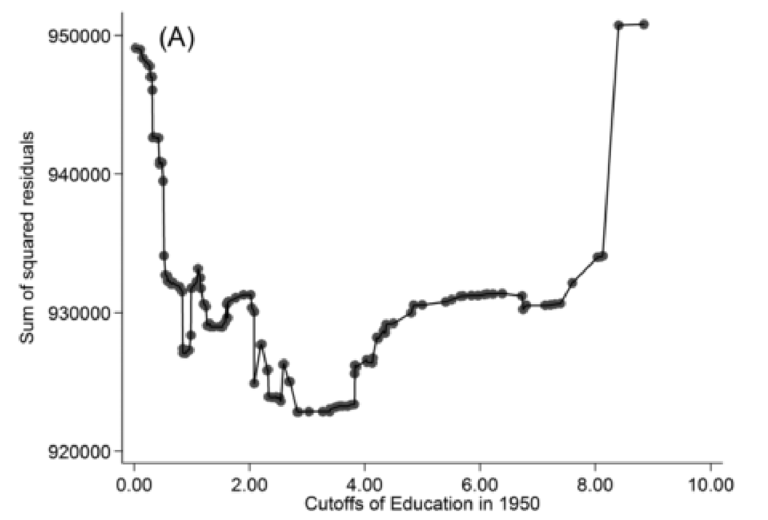
Relative importance: homogeneous equation (Shapley)

Panel B: Shapley values		
Base	0.011	0.009
Agriculture	0.423	0.321
Trade	0.033	0.025
Country FEs		0.222

Heterogeneity w.r.t. country "type"

- Idea: divide the world into early vs. late agglomerating countries
- Measures of historical structural transformation unfortunately are not available for all countries
- Solution 1: average years of schooling in the adult population in 1950 (Barro and Lee data)
- Solution 2: per capita GDO in 1950 (Madison Project)
- Solution 3: urbanization levels in 1950 (UN 2014)
- How to select the cutoff? Use the split the minimizes the unexplained variance (see next slide)

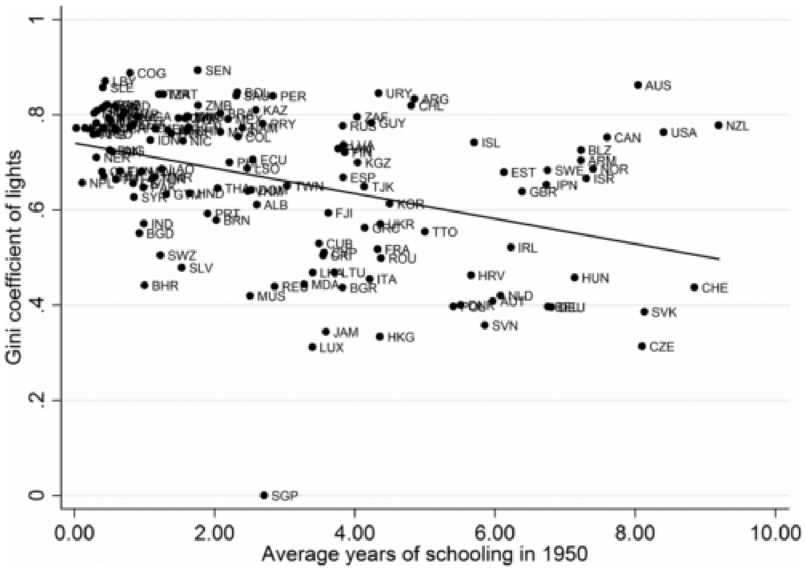
Cutoff for education



Differences across countries in relative importance

	Education		Urbanization		GDP per capita	
	High	Low	High	Low	High	Low
Countries	58	82	63	121	36	101
Observations	126,671	100,361	138,020	103,975	80,310	100,602
Panel A: R^2						
Full sample						
Base + FE	0.385	0.294	0.351	0.362	0.387	0.375
Agriculture + base + FE	0.653	0.452	0.614	0.511	0.644	0.521
Trade + base + FE	0.425	0.395	0.386	0.452	0.419	0.467
High - Low double differential	0.171		0.170		0.171	

Second finding: spatial inequality is higher in late agglomerating economies



GIS Data

Two types of GIS data

1. **Rasters:** images
   - in practice, grids with some resolution and values
   - some rasters have multiple bands: each band is a variable
   - e.g.: nightlights, MapBiomas, climate, DEM, constructed cost rasters
   - common file type: .tiff
2. **Shapefiles:** features
   - three types: points $\rightarrow$ lines $\rightarrow$ polygons
   - any georeferenced dataset can be turned into a points shapefile
   - e.g.: IBGE locations, roads, rivers, *setores censitários*, any administrative border
   - possible to convert between rasters and shapefiles
   - important: remember to simplify geometries of lines and polygons
   - common file type: .shp

How to work with GIS data?

- ArcGIS
- QGIS (and PyQGIS)
- Python
- R
- Stata (more recently)

Not every spatial data must be worked as GIS data

- GIS data is heavy, which increases processing cost
- A standard matrix dataset is much better to work on
  - do as little as you need on GIS data
- To calculate distances between points, we only need coordinates
- Know sphere geometry, especially Haversine formula:

$$\mathrm{hav}(\theta) = \sin^2(y_1 - y_0) + \cos(y_0)\cos(y_1)\sin^2(x_1 - x_0)$$

  - where θ is the ratio of distance between to points and the Earth radius $r \approx 6{,}378$ km

$$\text{distance} = 2r \arcsin\left(\sqrt{\mathrm{hav}(\theta)}\right)$$

Main data operations

- Merge or crop shapefiles or rasters
- Raster calculations: create new rasters from others
- Zonal statistics: assign raster values to polygons
- Join attributes by location: assign shapefile information
- Create buffers around lines or points
- Calculate distance between features (NNjoin)
- Calculate centroid and area of polygons

Pay attention to cartographic projections

- Standard projection: EPSG:4326, World Geodetic System 1984
  - coordinates on a plain
  - unit: degree
  - EPSG:4674 SIRGAS 2000 is very similar and common for IBGE data
- Distances and areas must be calculated in metric projections
  - some projections maintain area: Albers South America
  - all projections distort distances
  - for small areas (e.g. within city), use Universal Transverse Mercator (UTM) projections

Looking forward to this class!

bruno.barsanetti@fgv.br