GIS for Economists 1

Giorgio Chiovelli Sebastian Hohmann Tanner Regan

11/05/2023

Table of contents

Four aims

- Give introductions on:
  - *Geographic Information Systems* (GIS) methods, and how they are used in current economics research.
  - *Python* programming in general and the *GeoPandas* and related packages in particular.
- Show you how to:
  - "do GIS" for your own research through detailed examples of well published papers.
  - process large-scale satellite imagery in the cloud with Google Earth Engine.

We want to make this useful for you!

- You should install the software (Python / QGIS), so you can follow what we do in class.
- We will try to leave 10 minutes at the end of each hour for a break, and any concrete questions you may have.

What we will do in each session

1. (today) Overview of GIS applications in economics research, types of geospatial data, projections
2. (16/05/2023) Setting up Python, intro to Python, Pandas, Geopandas, simple examples.
3. (18/05/2023) GIS Basics. Replication of Michalopoulos (2012). "The Origins of Ethnolinguistic Diversity," American Economic Review, 102(4): 1508-1539.
4. (23/05/2023) Network analysis in GIS. Replication of Hornbeck and Donaldson. (2016). "Railroads and American Economic Growth: A Market Access Approach," Quarterly Journal of Economics (2016), 131(2): 799-858.
5. (25/05/2023) Selected topics: Installing / intro to QGIS, geocoding, making maps, georeferencing, and digitizing.
6. (30/05/2023) Intro to Google Earth Engine and Workshop Revision

All materials: Google $\boxed{\text{drive}}$ **Setup Guide**: here

Requirements for letter of participation and performance

1. Attendance at all six workshop lectures
2. Satisfactory submission of workshop project
   - The project will be assigned after the class on May 23
   - The project will be due on June 6th (one week after the last class)
3. Satisfactory score on multiple choice quiz
   - The quiz will be held online after the final class (May 30)

Introduction

Why use GIS in Economics Research? Felixibility in units of analysis!

Without GIS, spatial units of analysis are limited to

- countries
- administrative districts
  - (e.g. NUTS in Europe, counties in the US, etc.)
  - Can be very large, not consistent across countries
- survey locations
  - (e.g. villages, households, etc.)
  - Hard to scale, sampling problems

With GIS, the unit can be any level of spatial aggregation

- populate (spatially) administrative units
- non-administrative units
  - (macro) locations of ethnic groups, historical kingdoms, etc.
  - (micro) plots of land, individual buildings, etc.
- artificial units (gridcells, dyads, etc.)

More credible identification strategies.

- account for many geographic covariates
- create instruments
  - e.g. distance from: Mainz (spread of printing press) or ports (trade generally)
- conduct spatial RD-design
  - e.g. (macro) colonial borders in Africa, (micro) planned and parcelled urban land
- Explore level at which pattern uncovered prevails. Invariance to spatial aggregation?

Based on Stelios Michalopoulos's slides on GIS for Economic History

J. Vernon Henderson, Adam Storeygard David N. Weil (2012) “Measuring Economic Growth from Outer Space,” AER 102(2): 994–1028.

Motivation
Income / output is poorly measured, especially in developing countries and at small scales.

Contribution
Use data from NASA satellites capturing light of human settlements at night and develop method to aggregate these with official growth statistics.

Paper examples: satellite data

Night lights 1: "Pure measurement". Henderson, Storeygard and Weil (2012)

Night lights in 2008

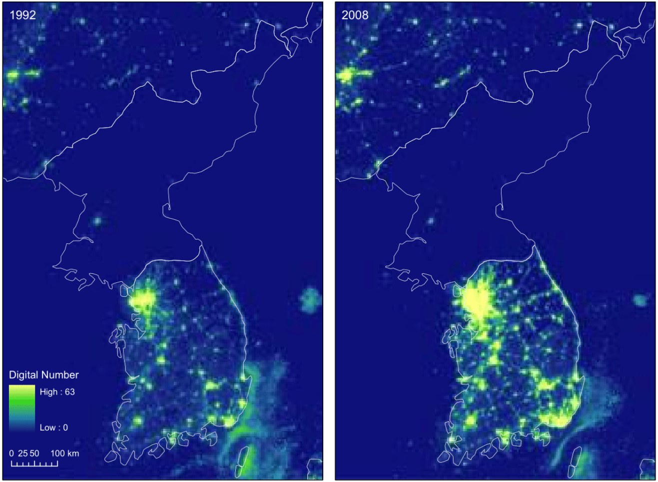
source:

https://eogdata.mines.edu/products/dmsp/

Paper examples: satellite data

Night lights 1: "Pure measurement". Henderson, Storeygard and Weil (2012)

Long-term growth: Korea

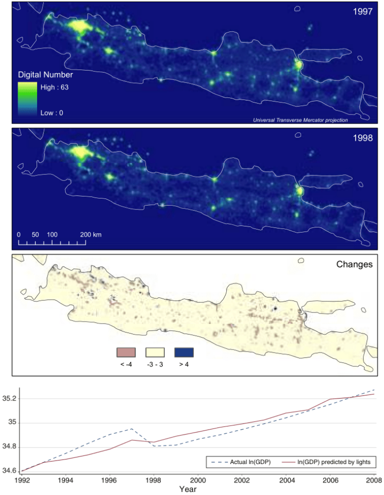


Universal Transverse Mercator projection

Paper examples: satellite data

Night lights 1: "Pure measurement". Henderson, Storeygard and Weil (2012)

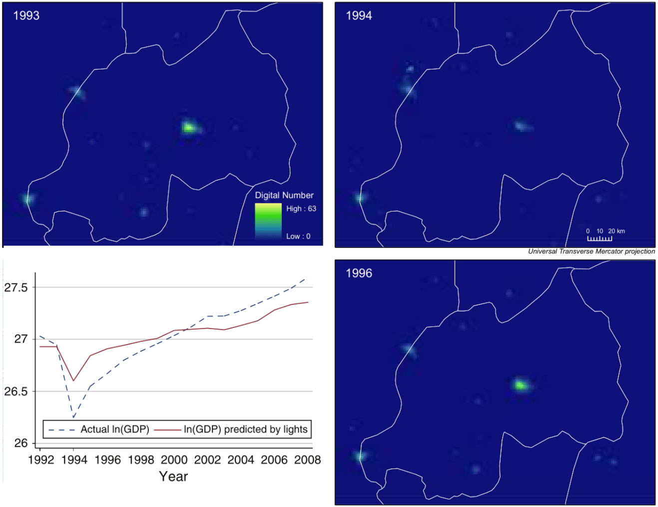
Short-term economic shocks: Indonesia 1997/1998



Paper examples: satellite data

Night lights 1: "Pure measurement". Henderson, Storeygard and Weil (2012)

The impact of war: Rwanda 1994



Paper examples: satellite data

Night lights 2: Institutions and development. Michalopoulos and Papaioannou (2014)

Elias Papaioannou and Stelios Michalopoulos (2014) “National Institutions and Sub-national Development in Africa,” QJE 129(1): 151-213.

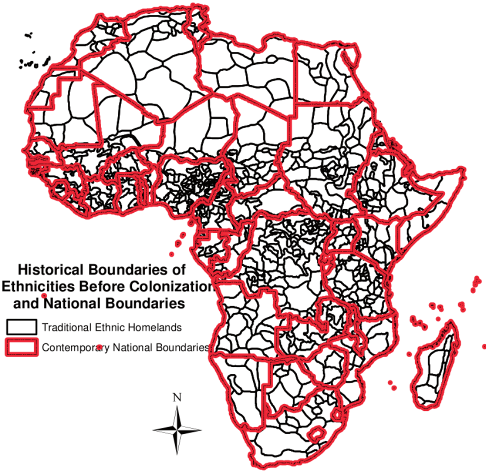
Motivation
Big literature on the role of institutions in economic growth (Acemoglu et al. (2001) etc.; La Porta et al (1997) etc.; Engerman and Sokoloff (2000) etc.) but identification challenge since institutions are endogenous.

Contribution
Combine Satellite data on night lights, old maps of ethnic groups ancestral homelands, and modern national borders to identify effect of national institutions on local development in Africa.

Paper examples: satellite data

Night lights 2: Institutions and development. Michalopoulos and Papaioannou (2014)

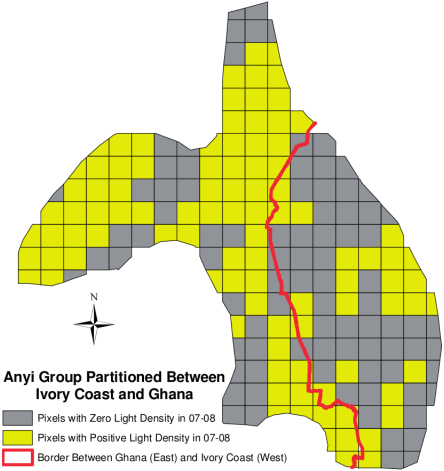
Pre-colonial ethnic and modern national boundaries



Paper examples: satellite data

Night lights 2: Institutions and development. Michalopoulos and Papaioannou (2014)

An example of a split ethnic group



Estimation (Linear probability model)

p = pixel

e = ethnicity

c = country

$$\mathbb{I}(pix\_lit_{p,e,c}) = \alpha_e + \beta \times inst\_qual_c + controls \tag{1}$$

Focussing on partitioned ethnicities only, including α_e kills the coefficient on institutional quality.

Roland Hodler and Paul Raschky (2014). "Regional Favoritism," QJE 129 (2): 995-1033.

Motivation
Large literature on favoritism from Bates (1974) onwards, but much of it anecdotal, case studies. Limited evidence across countries shows large heterogeneity across policy outcomes (Kramon and Posner (2013)).

Contribution
Systematically assess favoritism in large sample of countries: 38,427 sub-national regions from 126 countries, spanning democracies and autocracies. Exploit change in political leadership and information about birthplaces of countries' political leaders.

Paper examples: satellite data

Night lights 3: Favoritism. Hodler and Raschky (2014)

Mobutu in Zaire / DRC

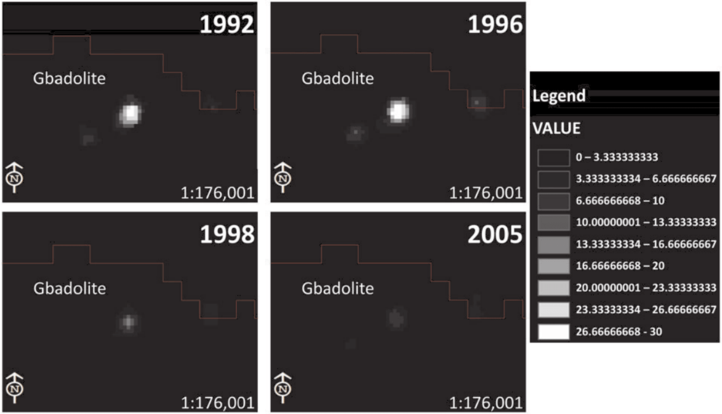


FIGURE I

Nightime Light Intensity in Gbadolite in 1992, 1996, 1998, and 2005

Mobutu Sese Seko was president of Zaire until 1997.

Paper examples: satellite data

Night lights 3: Favoritism. Hodler and Raschky (2014)

Mahinda in Sri Lanka

FIGURE II

Nighttime Light Intensity in Hambantota in 2003 and 2006

Mahinda Rajapaksa became prime minister of Sri Lanka in 2004 and president in 2005.

Yeh, Perez, Driscoll, Azzari, Tang, Lobell, Ermon and Burke (2020). “Using publicly available satellite imagery and deep learning to understand economic well-being in Africa,” Nature Communications, vol. 11, no. 2583.

Motivation

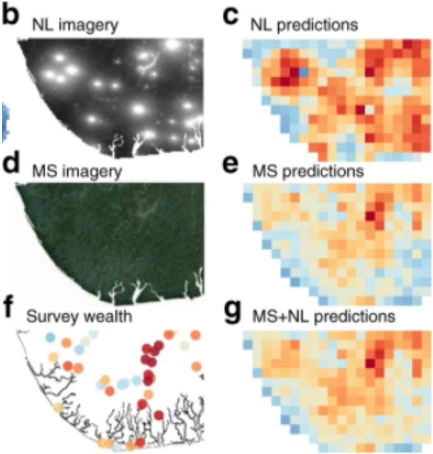
Measurements of wealth are important for research and policy, but often unavailable at a local level. Satellite images have become widely available, but turning them into meaningful data is difficult.

Contribution

Use night lights and high-resolution daytime (landsat) imagery. Use deep learning models to predict survey-based wealth across 20,000 African villages. Models can explain 70% of the variation in out-of-sample village wealth, and 50% of the changes.

Paper examples: satellite data
Measuring wealth: Yeh et al. (2020)

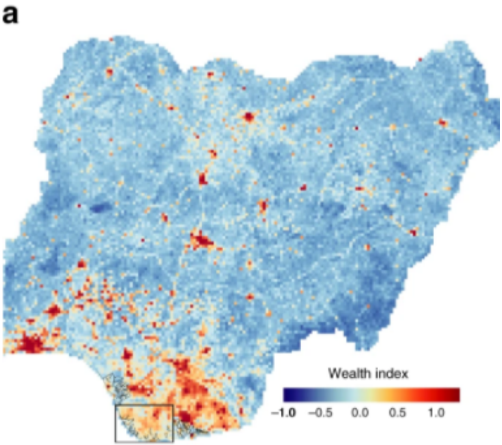
Input (b,d), training (f), and prediction (c,e,g) data.



Paper examples: satellite data

Measuring wealth: Yeh et al. (2020)

Nigeria predicted wealth for 6.7km^2 pixels.



Vernon Henderson, Tanner Regan and Anthony Venables (2021). "Building the City: From Slums to a Modern Metropolis," RES.

Motivation
Urbanisation is progressing rapidly in Africa, implying a huge demand for building volume. About two-thirds of a country's private capital stock are buildings. Yet, we know little about this investment process that reshapes cities.

Contribution
Model the building of a city with formal and slum construction technologies. Use very-high-resolution daytime imagery to trace individual buildings in 2003 and 2015 and measure redevelopment of built volume. Calculate the welfare loss due to delayed conversion of land.

Paper examples: satellite data

Urban development: Henderson et al. (2021)

Kibera, Nairobi VHR image for 2004 (left) and 2015 (right).

Paper examples: satellite data

Urban development: Henderson et al. (2021)

Kibera, Nairobi classified buildings for 2004 (left) and 2015 (right) - unchanged (blue), demolished (red), redeveloped (green).

Paper examples: satellite data

Ethnic patronage (housing investment): Marx et al. (2019)

Benjamin Marx, Thomas Stoker, and Tavneet Suri (2019). "There Is No Free House: Ethnic Patronage in a Kenyan Slum," AEJ:AE, 11(4): 36-70.

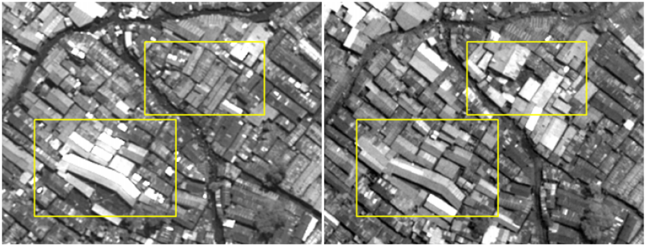
Motivation
Slums accommodate a growing fraction of the world's poor. At least 59 percent of sub-Saharan Africa's urban population are estimated to live in slums. Yet there remains little empirical evidence on the economic livelihoods of slum residents.

Contribution
Use very-high-resolution panchromatic (grayscale) images to proxy housing quality in a large slum. Provide evidence of ethnic patronage in the determination of rental prices and investments. Housing quality is lower when landlords have co-ethnic chiefs, and higher when residents have co-ethnic chiefs.

Paper examples: satellite data

Ethnic patronage (housing investment): Marx et al. (2019)

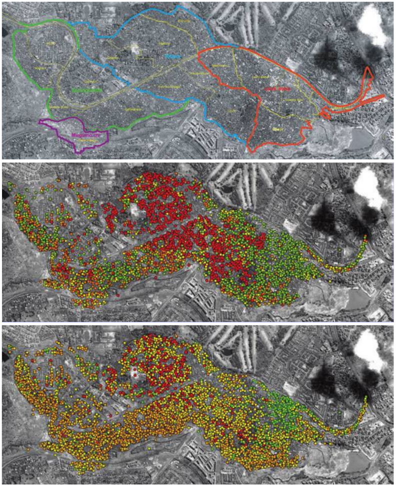
bright rooves=high quality housing - July 2009 (left) and August 2012 (right).



Paper examples: satellite data

Ethnic patronage (housing investment): Marx et al. (2019)

Kibera, Nairobi chiefs (top), landlords (middle), and households (bottom).



Paper examples: satellite data

Harnessing Imagery: Rolf et al. (2021)

Rolf, Proctor, Carleton, Bolliger, Shankar, Ishihara, Recht & Hsiang (2021). “A generalizable and accessible approach to machine learning with global satellite imagery,” Nature Communications, vol. 12, no. 4392.

Motivation
There is now an unbelievable amount of imagery data generated from satellites. Methods are being developed to classify and interpret this data, but often this is computationally prohibitive for economists or difficult to implement on case by case basis.

Contribution
This paper condenses the ‘contextual’ information from satellite images, into a smaller set of ‘features’. These features can then be used for many prediction exercises with simple (ridge regression, ‘lasso’) methods. Check out their data.

Yuyu Chen, Ginger Zhe Jin, Naresh Kumar, and Guang Shi (2013). "The Promise of Beijing: Evaluating the impact of the 2008 Olympic Games on air quality," JEEM 66 424-443.

Motivation

Pollution travels. Can drastic local policy interventions improve air quality?

Contribution

Show that measures to close roads and factories around the time of the 2008 Olympics improved air quality – both according to official statistics and to data from NASA satellites. Pollution quickly reverted to original levels after the games.

Pollution: Chen et al. (2013)

Change in pollution around the games from the paper, AOD data*

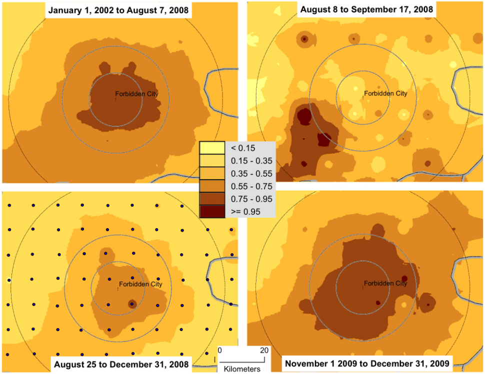


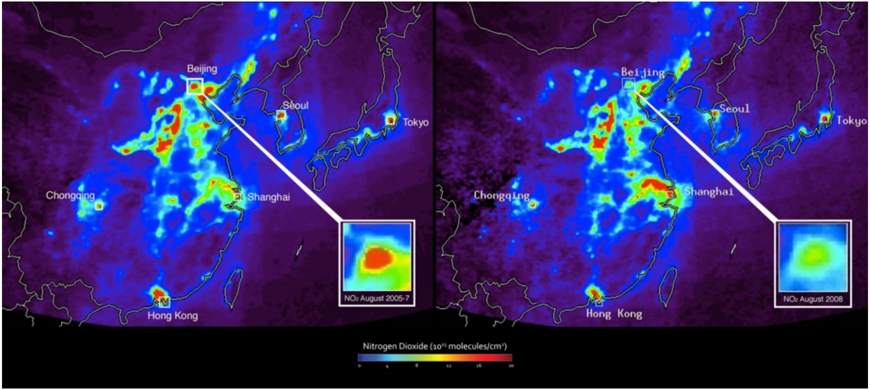
Fig. 7. Satellite-based AOD, corrected for meteorological conditions and spatiotemporal trends in and around Beijing.

* "Aerosol Optical Depth (AOD) captures the amount of radiation absorbed, reflected, and scattered due to the presence of solid and liquid particulates suspended in the atmospheric colum."

Paper examples: satellite data

Pollution: Chen et al. (2013)

NO_2 concentration before and during the games



Paper examples: satellite data

Deforestation: Burgess et al. (2012)

Robin Burgess, Matthew Hansen, Benjamin Olken, Peter Potapov, and Stefanie Sieber (2012). "The Political Economy of Deforestation in the Tropics," QJE 127 (4): 1707-1754.

Motivation

Tropical deforestation accounts for one fifth of greenhouse gas emissions. Much of it comes from illegal logging. What are the economic incentives driving this?

Contribution

Use satellite data on changes in forest cover to examine how local officials respond to changing incentives by allowing more or less logging.

Consistent with Cournot model: creation of new districts leads to higher deforestation and lower timber prices.

Paper examples: satellite data

Deforestation: Burgess et al. (2012)

Deforestation in Riau province

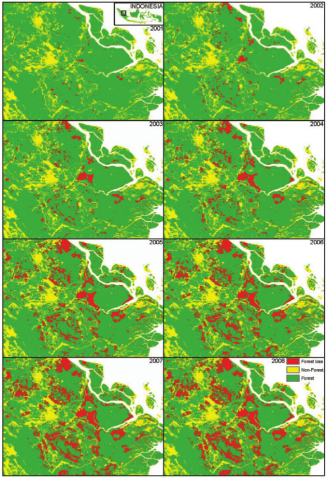


FIGURE I

Forest Cover Change in the Province of Riau, 2001–2008

Paper examples: satellite data

Deforestation: Burgess et al. (2012)

Deforestation in Indonesia 2001-2008

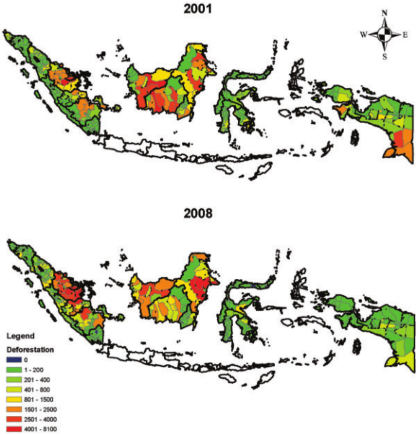


FIGURE II

District-Level Logging in Indonesia Using the 2008 District Boundaries, 2001 and 2008

Nathan Nunn (2008). "The Long-term Effects of Africa's Salve Trades," QJE 123(1): 139-176.

Motivation
Can Africa's underdevelopment be explained by the slave trade?

Contribution
Using archival research and GIS techniques, constructs estimates of slavery-intensity at the country-level. Uses these estimates, together with distance-based IV (also constructed with GIS) to show that slavery had negative impact on development among African countries.

Paper examples: old maps

The impact of slavery: Nunn (2008)

Constructing slavery-intensity step 1

TABLE I
SLAVE ETHNICITY DATA FOR THE TRANS-ATLANTIC SLAVE TRADE

Location	Years	Num. ethnic.	Num. obs.	Record type
Valencia, Spain	1482–1516	77	2,675	Crown records
Puebla, Mexico	1540–1556	14	115	Notarial records
Dominican Republic	1547–1591	26	22	Records of sale
Peru	1548–1560	16	202	Records of sale
Mexico	1549	12	80	Plantation accounts
Peru	1560–1650	30	6,754	Notarial records
Lima, Peru	1583–1589	15	288	Baptism records
Colombia	1589–1607	9	19	Various records
Mexico	1600–1699	28	102	Records of sale
Dominican Republic	1610–1696	33	55	Government records
Chile	1615	6	141	Sales records
Lima, Peru	1630–1702	33	409	Parish records
Peru (Rural)	1632	25	307	Parish records
Lima, Peru	1640–1680	33	936	Marriage records
Colombia	1635–1695	6	17	Slave inventories
Guyane (French Guiana)	1690	12	69	Plantation records
Colombia	1716–1725	33	59	Government records
French Louisiana	1717–1769	23	223	Notarial records
Dominican Republic	1717–1827	11	15	Government records
South Carolina	1732–1775	35	681	Runaway notices
Colombia	1738–1778	11	100	Various records
Spanish Louisiana	1770–1803	79	6,615	Notarial records
St. Dominique (Haiti)	1771–1791	25	5,413	Sugar plantations
Bahia, Brazil	1775–1815	14	581	Slave lists
St. Dominique (Haiti)	1778–1791	36	1,280	Coffee plantations
Guadeloupe	1788	8	45	Newspaper reports
St. Dominique (Haiti)	1788–1790	21	1,297	Fugitive slave lists
Cuba	1791–1840	59	3,093	Slave registers
St. Dominique (Haiti)	1796–1797	56	5,632	Plantation inventories
American Louisiana	1804–1820	62	223	Notarial records
Salvador, Brazil	1808–1842	6	456	Records of manumission
Trinidad	1813	100	12,460	Slave registers
St. Lucia	1815	62	2,333	Slave registers
Bahia, Brazil	1816–1850	27	2,666	Slave lists
St. Kitts	1817	48	2,887	Slave registers
Senegal	1818	17	80	Captured slave ship
Berbice (Guyana)	1819	66	1,127	Slave registers
Salvador, Brazil	1819–1836	12	871	Manumission certificates
Salvador, Brazil	1820–1835	11	1,106	Probate records
Sierra Leone	1821–1824	68	605	Child registers
Rio de Janeiro, Brazil	1826–1837	31	772	Prison records
Anguilla	1827	7	51	Slave registers
Rio de Janeiro, Brazil	1830–1852	190	2,921	Free africans' records
Rio de Janeiro, Brazil	1833–1849	35	476	Death certificates

TABLE I
(CONTINUED)

Location	Years	Num. ethnic.	Num. obs.	Record type
Salvador, Brazil	1835	13	275	Court records
Salvador, Brazil	1838–1848	7	202	Slave registers
St. Louis/Goree, Senegal	1843–1848	21	189	Emancipated slaves
Bakel, Senegal	1846	16	73	Sales records
d'Agoué, Benin	1846–1885	11	70	Church records
Sierra Leone	1848	132	12,425	Linguistic and British census
Salvador, Brazil	1851–1884	8	363	Records of manumission
Salvador, Brazil	1852–1888	7	269	Slave registers
Cape Verde	1856	32	314	Slave census
Kikoneh Island, Sierra Leone	1896–1897	11	185	Fugitive slave records

Paper examples: old maps

The impact of slavery: Nunn (2008)

The Murdock (1959) tribal map as digitized by Nunn (2008)

Also used by Michalopoulos and Papaioannou (ECMA 2013, QJE 2014, AER 2016).

Constructing slavery-intensity step 2

FIGURE II
Ethnic Boundaries Defined by Murdock (1959) and Modern Political Boundaries

Paper examples: network

Transport Costs and Economic Development: Storeygard (2016)

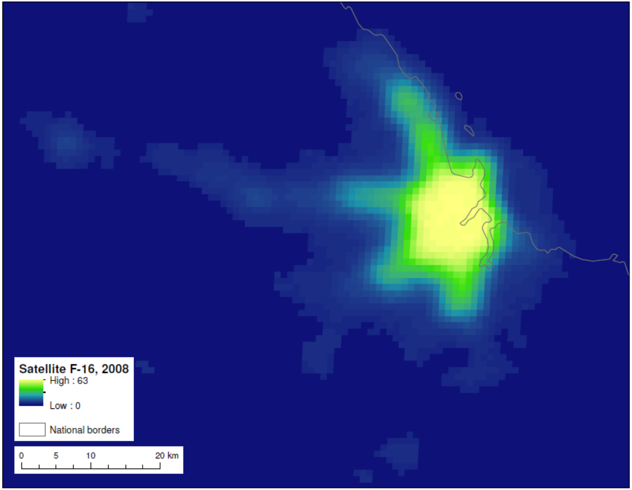
Storeygard, Adam (2016). "Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa" RES 123(1): 139-176.

Motivation
How inter-city connectivity determines the income of sub-Saharan African cities.

Contribution
Using road quality/type, lights, and oil price shock, estimates how growth is propagated far away from the city port. Negative elasticity of city economic activities wrt transport costs. Effect is heterogeneous in road quality (paved/unpaved).

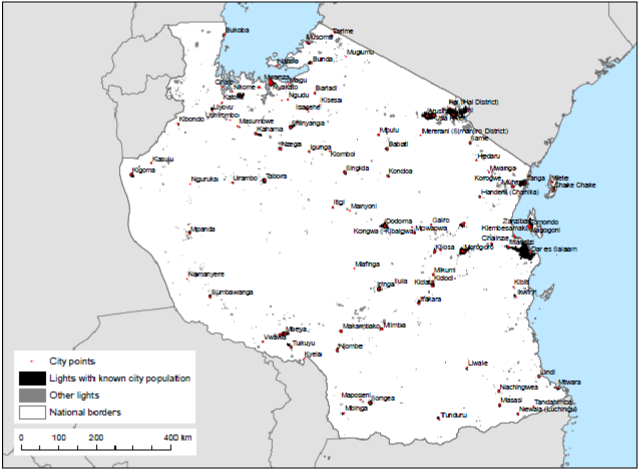
Paper examples: network

Luminosity (2008) in Dar es Salaam



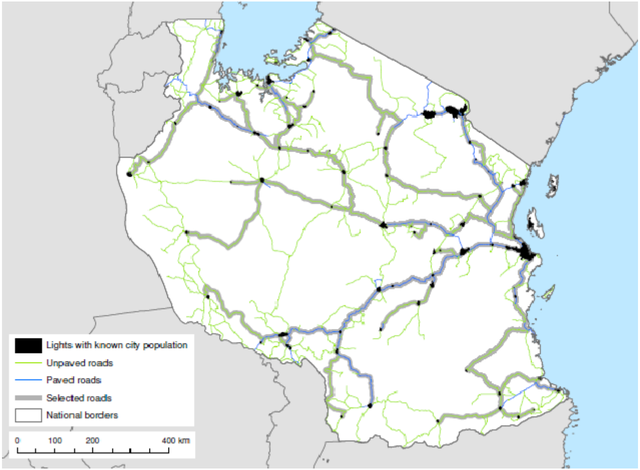
Paper examples: network

City Distribution in Tanzania



Paper examples: network

Roads connection to Dar es Salaam (Tanzania)



Paper examples: network

Trafficking Networks and the Mexican Drug War: Dell (2016)

Dell, Melissa (2016). "Trafficking Networks and the Mexican Drug War" AER

Motivation

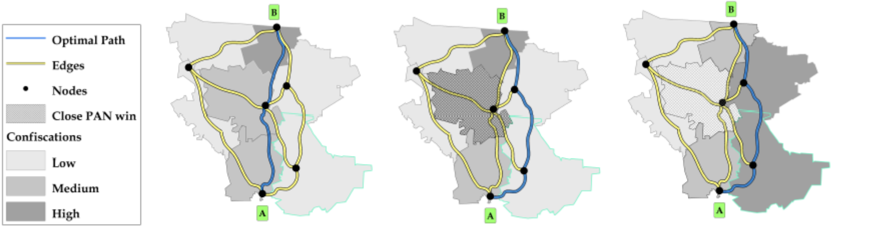
How crime repression policy affects spatial pattern of violence and crime .

Contribution

Using close elections, disaggregated crime data, and road network, to identify how crime reacts to crackdown policy. Quantification of externalities and unintended consequences of crime repression.

Paper examples: network

Spillover Methodology



(a) Legend

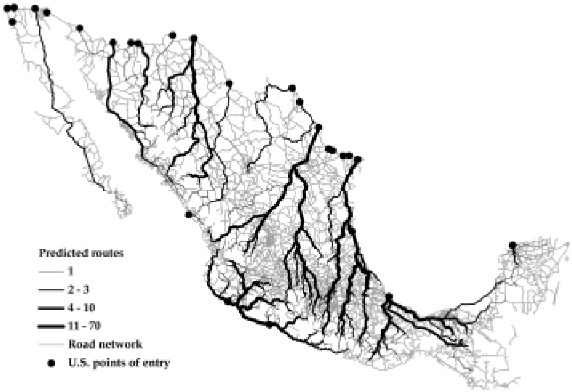
(b) Basic environment

(c) Close PAN victory increases costs; routes change

(d) Do confiscation patterns also change?

Paper examples: network

Roads Network and Predicted Trafficking Routes



Donaldson, Dave and Hornbeck, Richard (2015). "Railroads and American Economic Growth: A "Market Access" Approach" QJE

Motivation
Evaluate the role of railway construction for US economic growth.

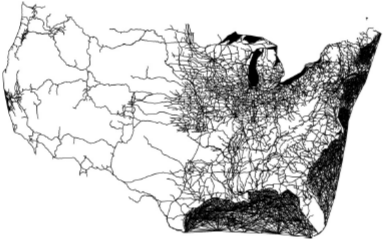
Contribution
Combine transportation network over time and census data to show that expansion of the railway network fostered both local and aggregated economic growth. Theory-based application of intra-country trade model (Eaton and Kortum, 2002) in a reduced form framework to quantify the spillover effect of infrastructure project.
Railways expansion fostered US economic growth. Aggregate effects are considerably larger than local effects (due to higher "market access").

Paper examples: network

Railway expansion 1870-1890

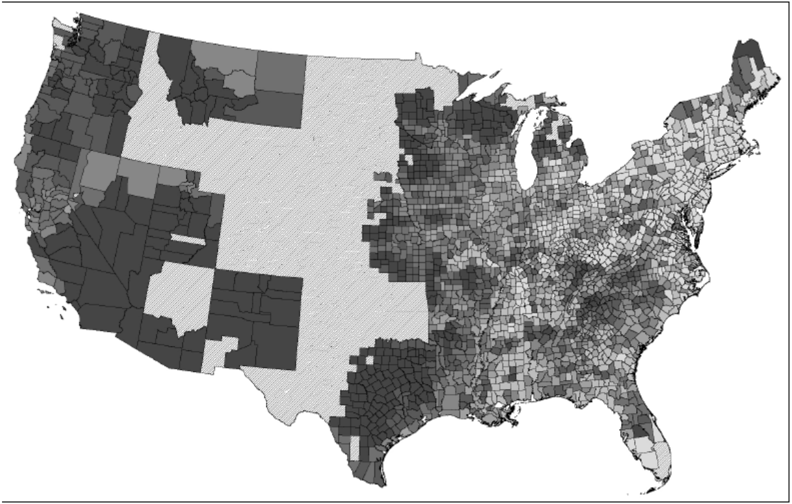
C. Natural Waterways, Canals, and 1870 Railroads

D. Natural Waterways, Canals, and 1890 Railroads



Paper examples: network

Change in Market Access 1870-1890



Paper examples: network

Archeology!: Barjamovic, Chaney, Cosar and Hortacsu (2019)

Gojko Barjamovic, Thomas Chaney, Kerem Cosar, and Ali Hortacsu (2019). "Trade, Merchants and the Lost Cities of the Bronze Age" QJE

Motivation
Test gravity model of trade, find lost cities!

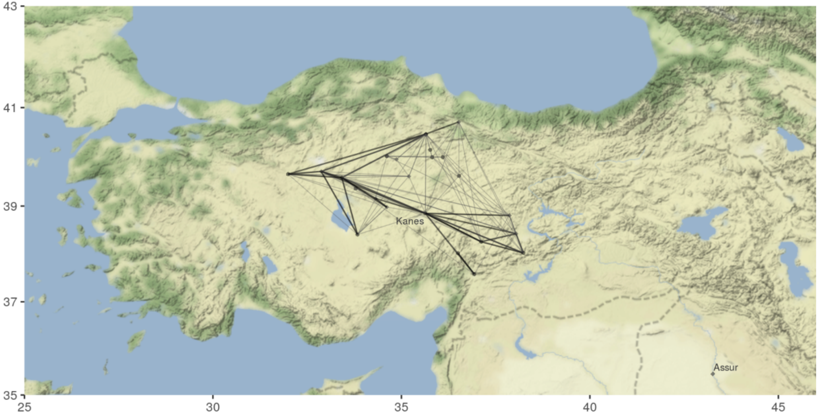
Contribution

- extract information on trade linkages between Bronze-age Assyrian cities: how many transactions between cities?
- specify structural gravity model of trade that can be estimated purely with: (1) trade flows between cities, (2) geographic distance between cities
- can "invert" the gravity model to pinpoint the likely locations of lost cities!

Paper examples: network

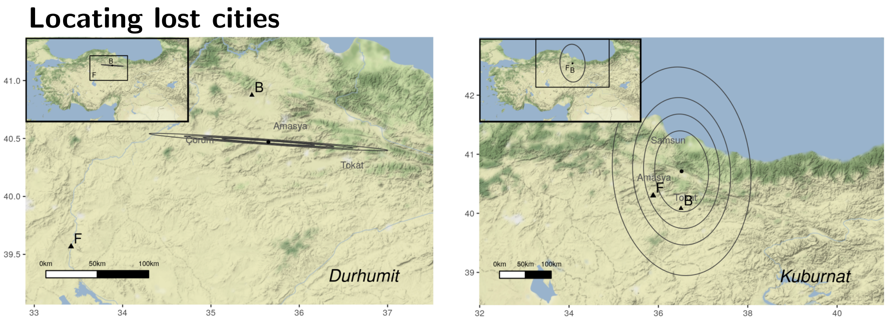
Archeology!: Barjamovic, Chaney, Cosar and Hortacsu (2019)

Ancient cities and trade routes



Paper examples: network

Archeology!: Barjamovic, Chaney, Cosar and Hortacsu (2019)



"dot" = point estimate of lost city location, "B/F" = Archeologist estimates of city locations

Stelios Michalopoulos and Elias Papaioannou (2016). "The Long-Run Effects of the Scramble for Africa," AER 106(7). 1802-1848

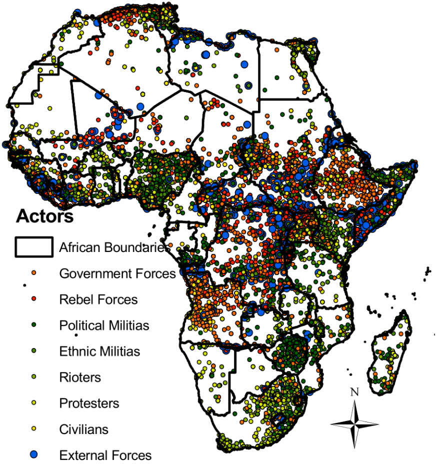
Motivation

Did the arbitrary partitioning of ethnic groups during the "Scramble for Africa" in the 19th century create a long-lasting legacy of violent conflict?

Contribution

Combine the digitized Murdock (1959) map with modern political boundaries and geo-referenced data on conflict events to show that ethnicities split between two modern states experience more violent conflict (incidence, severity, and duration). Identification credible since drawing of colonial borders "quasi-random" with respect to existing ethnic groups.

Conflict events as reported by ACLED*



* `http://www.acleddata.com/`

Paper examples: conflict

Conflict 2: Propaganda. Yanagizawa-Drott (2014)

David Yanagizawa-Drott (2014). "Propaganda and Conflict: Evidence from the Rwandan Genocide," QJE 129(4). 1947-1994

Motivation

Rwandan genocide (1994) killed 800,000-1,000,000 people. Quickest mass extermination in history. Prominent role of "hate radio" in mobilizing killers. → Can mass media be used to incite violence against civilians?

Contribution

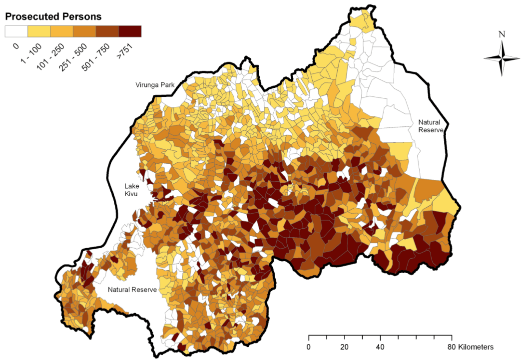
Exploits quasi-random variation in hate radio signal strength (combination of transmitter location and Rwanda's mountainous terrain) to show that broadcasts increased participation in killings by locals. Around 10% of deaths can be attributed to radio*.

* For similar identification strategy exploiting variation in signal strength see Enikolopov, Ruben, Maria Petrova and Ekaterina Zhuravskaya (2011) "Media and Political Persuasion: Evidence from Russia". AER, 101(7): 3253-85.

Paper examples: conflict

Conflict 2: Propaganda. Yanagizawa-Drott (2014)

Genocide violence in villages

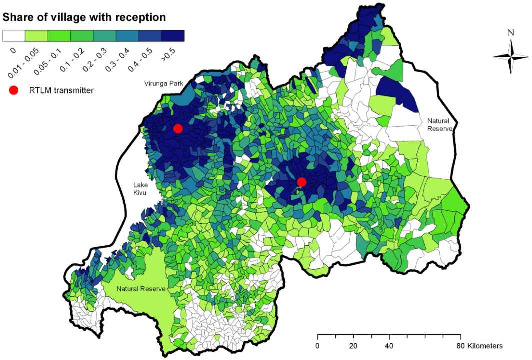


The categories represent the total number of prosecuted persons in the village (sum of militia and individual violence). White areas are missing data, either because of geography, such as parks and natural reserves, or villages that lack data in the sample.

Paper examples: conflict

Conflict 2: Propaganda. Yanagizawa-Drott (2014)

RTLM radio coverage



The figure shows the radio coverage in villages (share of village area with sufficient radio reception) based on the Longley-Rice propagation model. Source: Author's calculations in ArcGIS using the Longley-Rice Propagation Model.

Paper examples: conflict

Conflict 3: The role of militias as motivators. Rogall (2021)

Thorsten Rogall, (2021). "Mobilizing the masses for genocide," AER 111(1). 41-72

Motivation

Did organized militias matter in as organizers and motivators for civilians killing other civilians?

Contribution

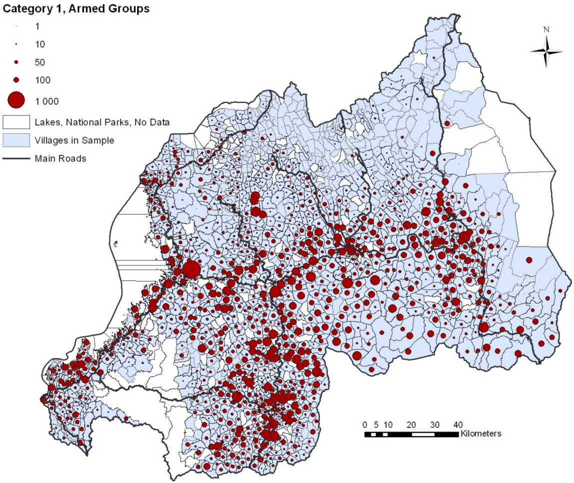
Exploits quasi-random variation in transport costs of militias in accessing villages created by different amounts of rain falling on unpaved roads during the genocide.

Main Finding: Militia presence did increase killing rates. One additional militiaman resulted in 13 additional deaths.

Paper examples: conflict

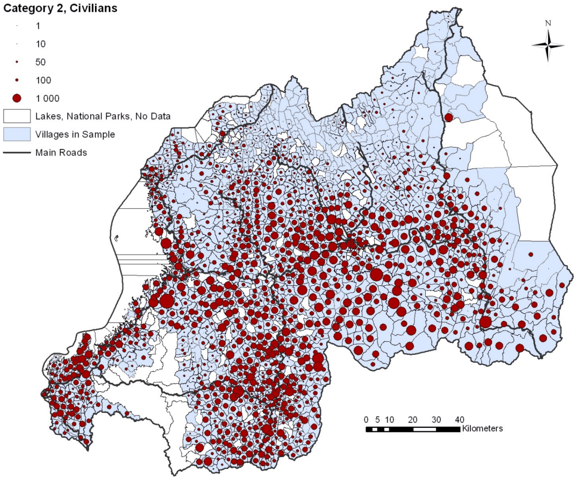
Conflict 3: The role of militias as motivators. Rogall (2021)

Violence by armed groups



Conflict 3: The role of militias as motivators. Rogall (2021)

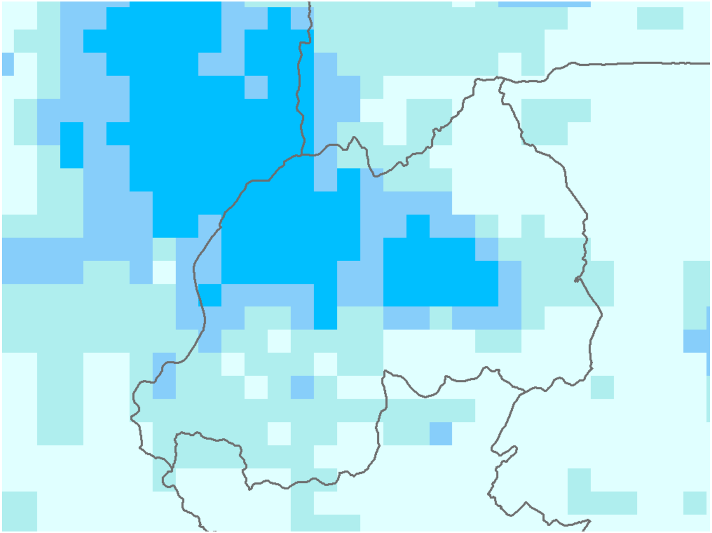
Violence by civilians



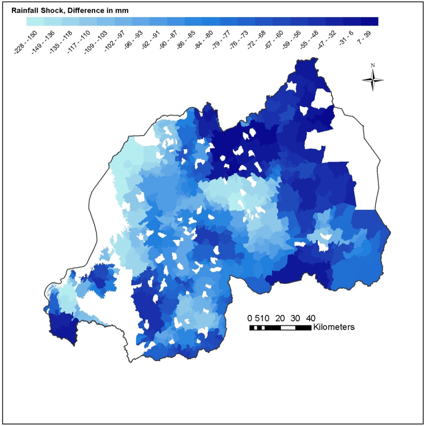
Paper examples: conflict

Conflict 3: The role of militias as motivators. Rogall (2021)

Rainfall Data Example - NOAA (NASA)



Detailed rainfall data



Note: This map shows rainfall along the way between main road and village during the period of the genocide in 1994 for each village, subtracting rainfall between main road and village during the 100 calendar days of the genocide of an average year (years 1984-1993). White areas are either national parks, Lake Kivu or villages not in the sample.

Paper examples: conflict

Conflict 3: The role of militias as motivators. Rogall (2021)

Instrument construction

Instrument: Interaction of the length of the red line and amount of rain falling on the area of the blue rectangle during the period of the genocide.

Rohini Pande and Esther Duflo (2007). "Dams," QJE, 122(2) :601-646.

Motivation
Construction of dams for irrigation and power generation is (a) widespread (45,000 large dams, 1/2 of rivers in the world have one), (b) thought to lead to development and poverty reduction.

Contribution
Overcome endogeneity problem of dam construction using river gradient as instrument for dam suitability.
In panel of Indian districts, find distributional effects:
downstream districts benefit (output mean $\nearrow$, output volatility, poverty $\searrow$),
dam districts lose (output mean $\rightarrow$, output volatility, poverty $\nearrow$).

Paper examples: other applications

Other applications 1: Topography as an instrument. Pande and Duflo (2007)

Dams by District in 1970

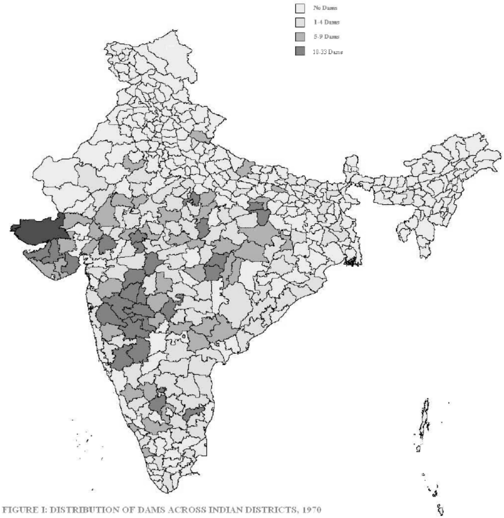


FIGURE I: DISTRIBUTION OF DAMS ACROSS INDIAN DISTRICTS, 1970

Paper examples: other applications

Other applications 1: Topography as an instrument. Pande and Duflo (2007)

Dams by District in 1999

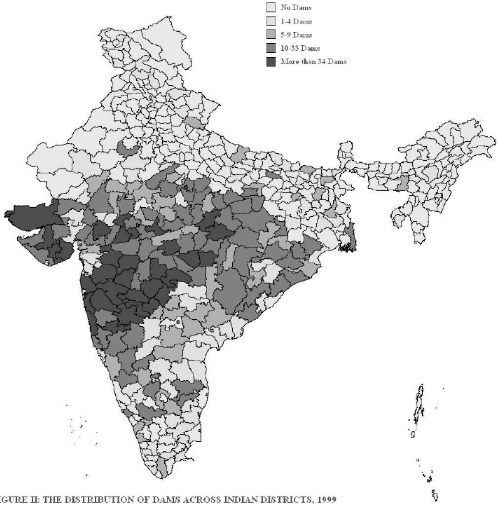


FIGURE II: THE DISTRIBUTION OF DAMS ACROSS INDIAN DISTRICTS, 1999

Paper examples: other applications

Other applications 1: Topography as an instrument. Pande and Duflo (2007)

Average River Gradient (Percentage) by District

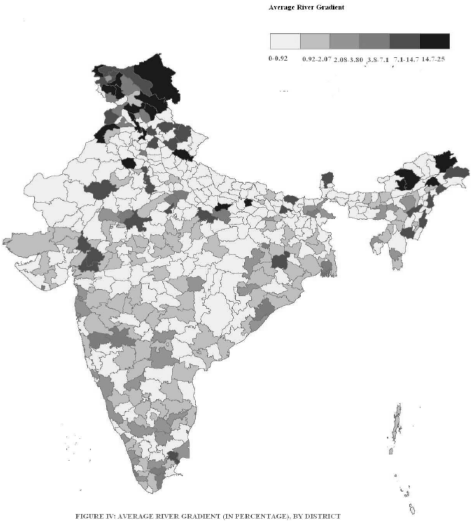


FIGURE IV: AVERAGE RIVER GRADIENT (IN PERCENTAGE), BY DISTRICT

Other examples of topography instruments

- Dinkelman (2011): land gradient as IV for electrification.
- Qian (2008): land gradient as IV for tea production.
- Olken (2009): signal strength (blocked by mountains) as IV for TV availability.
- Durante, Pinotti, and Tesei (2017): similar strategy as Olken (2009)
- Yanagizawa-Drott (2014): Same idea (see above).
- Lipscomb et al. (2013): Topography measures (river gradient etc.) as IV for hydropower.

Melissa Dell (2010). "The Persistent Effects of Peru's Mining Mita," ECMA. 78 (6) :1863-1903.

Motivation
Are there long-run effects of historical institutions (the forced-labor system of Peru's mining "Mita") on today's development? If so, what are the channels?

Contribution
One of the first to use spatial RDD to cleanly identify effect. Finds economically meaningful effects (25% lower household consumption, 6% increase in childhood stunting). Speaks to channels: "Mita districts historically had fewer large landowners and lower educational attainment. Today, they are less integrated into road networks and their residents are substantially more likely to be subsistence farmers." (persistence through land tenure, public goods, market participation)

Paper examples: other applications

Other applications 2: Spatial RDD. Dell (2010)

Mita boundary

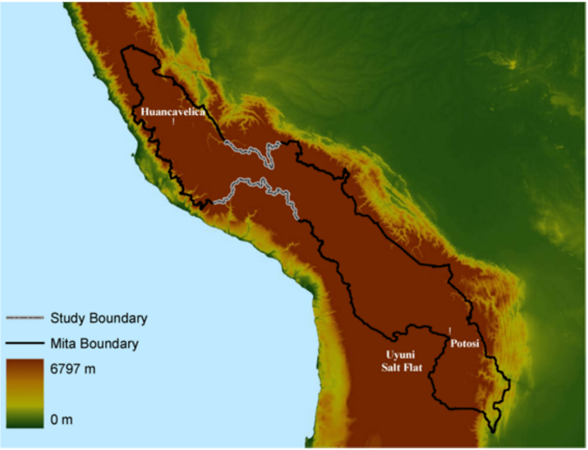


FIGURE 1.—The *mita* boundary is in black and the study boundary in light gray. Districts falling inside the contiguous area formed by the *mita* boundary contributed to the *mita*. Elevation is shown in the background.

Paper examples: other applications

Other applications 3: Urban spatial equilibrium. Ahlfeldt et al. (2015)

Ahlfeldt, Redding, Sturm, and Worlf (2015). "The Economics of Density: Evidence from the Berlin Wall," ECMA. 83 (6):2127–2189.

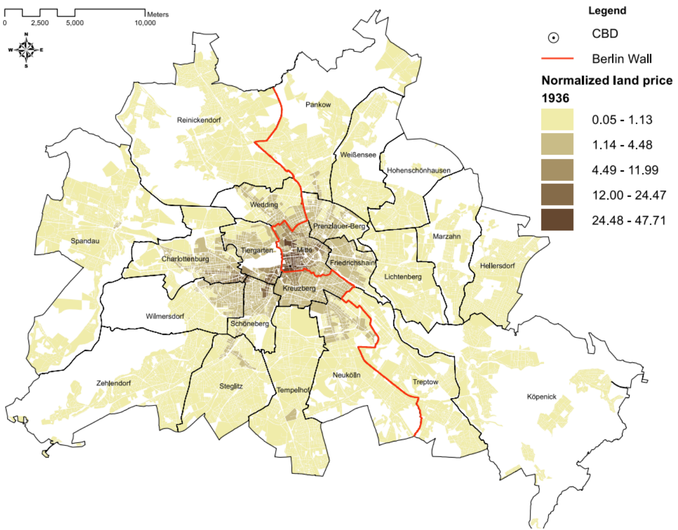
Motivation
Economic activity is highly unevenly distributed across space (Manhattan in New York and the Square Mile in London). Theoretically, agglomeration and dispersion forces underlie these patterns. But, distinguishing these forces from variation in locational fundamentals is difficult empirically.

Contribution
This paper was the first to develop an empirically tractable quantitative model of internal city structure that features endogenous agglomeration and dispersion forces. Use data on thousands of city blocks in Berlin for 1936, 1986, and 2006 and exogenous variation from the city's division and reunification.

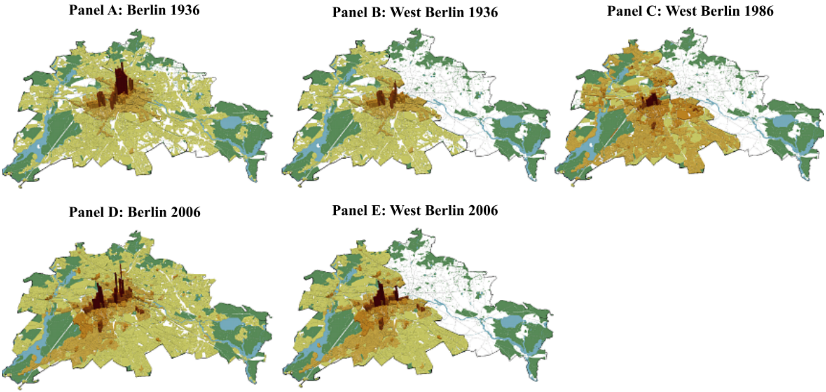
Paper examples: other applications

Other applications 3: Urban spatial equilibrium. Ahlfeldt et al. (2015)

Berlin 1936 land prices and the Wall



Berlin (east and west) evolution of land price - before, during, and after division



We will cover two things to start with GIS

- Types of geographic data
- Coordinate systems and projections

Geographic data comes in a vast number of formats.

For 99.9% of all applications you will only need two:

- **Feature** (vector) data, files end in *.shp* (shapefiles)
- **Raster** (cell) data, files (typically) end in *.tif*, but other formats are also common.

It is useful to distinguish three types of feature data

- polygon features
- polyline features
- point features

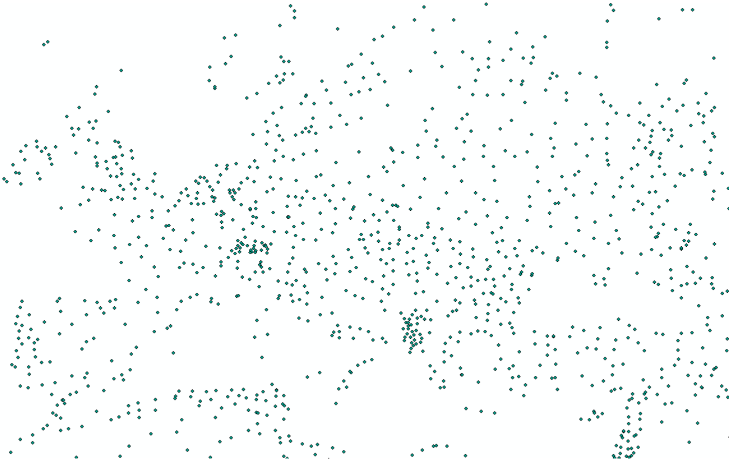
Countries can be treated as polygon features

Rivers as polylines

Introduction to GIS

Data types – point features

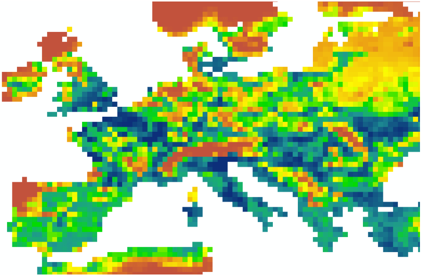
Cities as points



Introduction to GIS

Data types – raster data

Agricultural suitability in raster format



Each cell in the raster has a value

- cells can be empty
- data does not need to be integer (can be float)

Why this matters

- Geospatial data are data with geographic identifiers attached to each data point.
- To view, manipulate, and make calculations with geographic data, the identifiers need to be referenced with respect to a coordinate system.
- Each coordinate system represents the Earth's sphere (or a part of it) in two dimensions.
- Results of calculations depend on the coordinate system used. Using the correct one is crucial!

There are two basic types of coordinate systems

- Geographic coordinate systems.
- Projected coordinate systems.

Based on Melissa Dell's GIS notes.

Geographic coordinate systems represent locations in spherical coordinates .

It is standard to write information for geographic coordinates systems in two ways

- **degrees, minutes, seconds**
- e.g. London is 51° 30' 26" N, 0° 7' 39" W
- **decimal degrees**
- London is 51.5072, -0.1275
- 30 arc minutes and 26 arc seconds $= 30 \times 60 + 26 = 1826$ arc seconds $= \frac{1826}{3600} = 0.5072$ decimal degrees.

The standard geographic coordinate system is WGS 1984 ("World Geodetic System"). It is the only one ever used by economists.

For distances between two points along the Earth's surface, can use **geodesic distance formula**: need only $((lat_1, lon_1), (lat_2, lon_2))$.

Problem

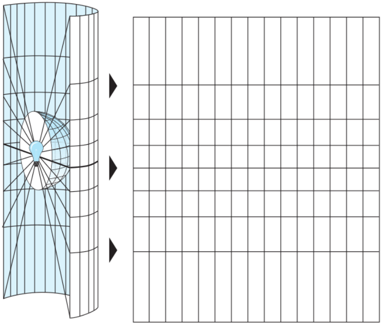
- Geographic coordinate systems are useless to calculate areas and distances along lines.
- Geographic coordinate systems distort lengths and areas: try to wrap a map around an orange!
- 1° latitude is 110.6km at the equator, 111.7km at the poles
- 1° longitude is 111.3km at the equator, 55.8km at 60° N/S

Mercator distortion example

To make accurate area and (some) distance calculations, need to **project** the data from a geographic coordinate system.

One can think of shining a light, placed on the centre of the earth, through the earth surface, and casting a shadow on a projection surface of a certain shape.

Projecting onto a cylindrical projection surface



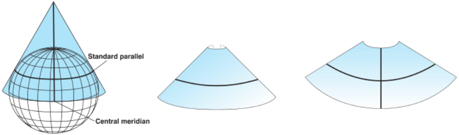
note stretching of data near the poles

Useful three-way classification of simple projections

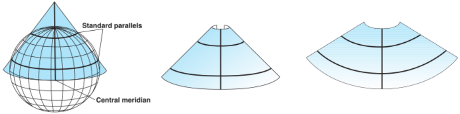
- conic
- cylindrical
- planar

Conic projections

Tangent conic projection, one standard parallel

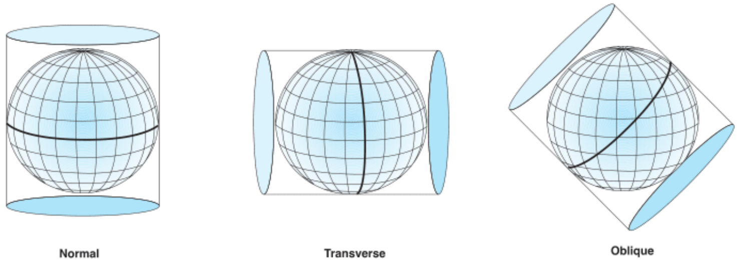


Secant conic projection, two standard parallels



Cylindrical projections

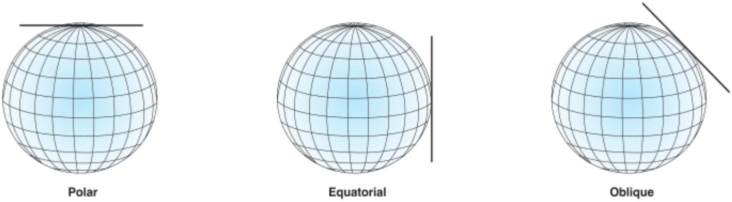
Cylindrical projections



Mercator projection is cylindrical with equator as line of tangency.

Planar projections

Planar projections with different point of tangency (*aspect*)



WGS 1984

- for distance between two points

UTM: divides surface of the earth into many regions, each gets its own projection

- distance/surface area in small regions
- length of polylines

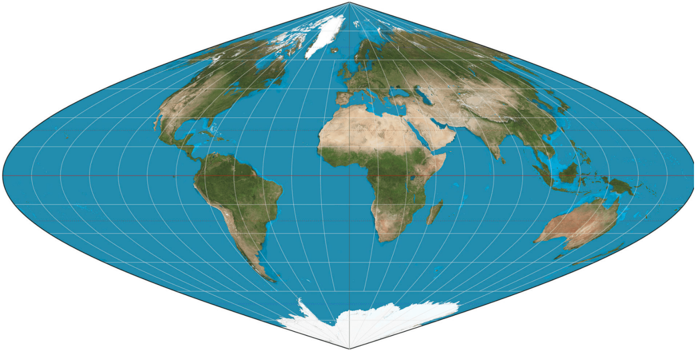
Any equal area projection

- surface area in large regions

This and the following three slides are based on Masayuki Kudamatsu's GIS course

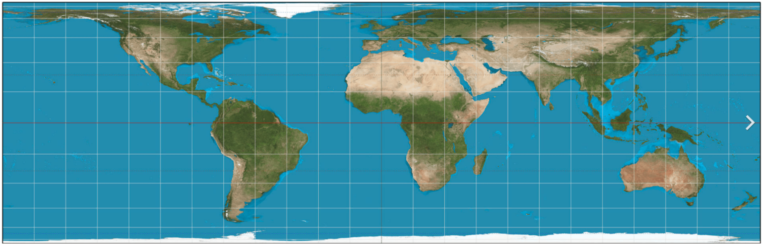
Differ just in how the world is shown

Sinusoidal projection



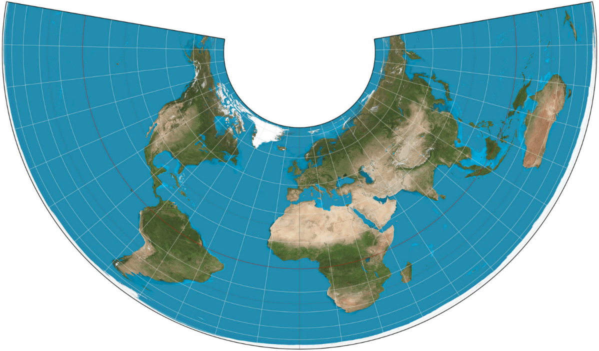
Differ just in how the world is shown

Lambert cylindrical equal area projection

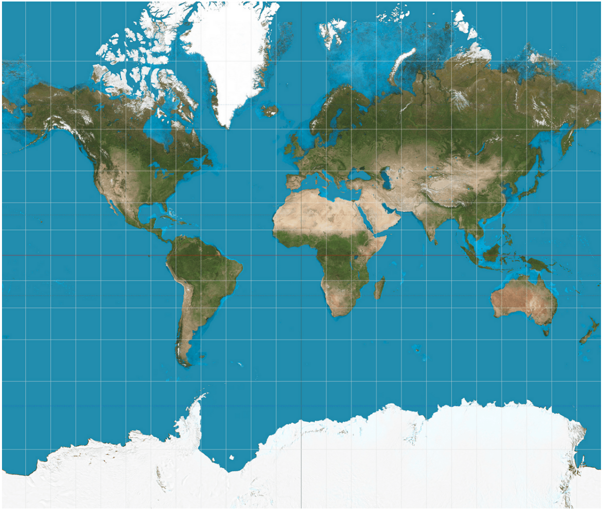


Differ just in how the world is shown

Albers Equal area conic projection



The world in Mercator projection



Appendix

The actual sizes of Greenland and Australia

nice *The Economist* article on map projections:

https://www.economist.com/graphic-detail/2016/12/27/misleading-maps-and-problematic-projections

back