MIT 14.582 PhD International Economics II — Lectures 19-20: Economic Geography and Urban Economics (Path Dependence) —

Dave Donaldson

Spring 2025

Plan for Today's Lecture

- Broad goal: different approaches to *path dependence* in economic geography settings
- Papers that look for direct evidence of path dependence:
  1. WWII bombing: Davis and Weinstein (2002) and Davis and Weinstein (2008)
  2. Portage: Bleakley and Lin (2012)
- Quantitative calculations about the importance of path dependence in the case of the US, 1850-2000: Allen and Donaldson (2020)

Plan for Today's Lecture

- Broad goal: different approaches to path dependence in economic geography settings
- **Papers that look for direct evidence of path dependence:**
  1. **WWII bombing: Davis and Weinstein (2002) and Davis and Weinstein (2008)**
  2. Portage: Bleakley and Lin (2012)
- Quantitative calculations about the importance of path dependence in the US: Allen and Donaldson (2020)

Path Dependence and Economic Geography

- As we've seen so far, evidence for agglomeration economies seems strong
- Long theoretical tradition highlights implications:
  - Potential for multiple equilibria in static models
  - Potential for *path dependence*: where "initial conditions" (whenever you call "initial"), and hence long-redundant shocks, still matter for outcomes today in dynamic models because they influence the steady-state that the economy converges to
  - Potential for policies to promote movement to better steady-state
- But is path dependence actually empirically consequential?
  - That is, given a particular setting, should we expect PD to occur?
  - If so, "history matters". But does history matter much for, e.g., the location of economic activity, and/or the total amount of economic activity (i.e. welfare)?

Davis and Weinstein (AER, 2002)

- DW (2002) ask how regions/cities' population levels responded to the shock of WWII bombing in Japan—basic idea is that this is a shock to "initial conditions" (c. 1946) and we can ask whether this has long-lived consequences
- Their findings are surprising—and yet have been replicated in many other settings:
  - Germany (WWII): Brakman, Garretsen and Schramm (2004)
  - Vietnam (Vietnam war): Miguel and Roland (2011)
  - ...
- Davis and Weinstein (J Reg. Sci., 2008) extend the analysis in DW (2002) to the case of the fate of industry-locations.
  - This is doubly interesting as it is plausible that industrial activity is mobile across space (i.e. might be happy to resettle in a new "equilibrium") in ways that people are not.

DW (2002): Convergence (non-path dependence?) for 2 cities

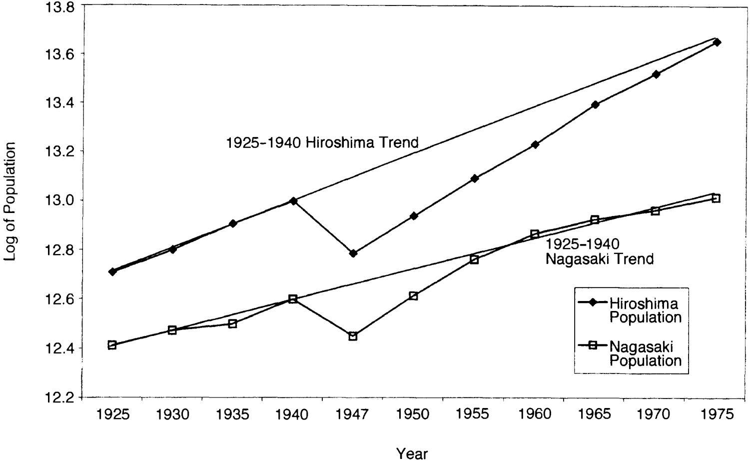


FIGURE 2. POPULATION GROWTH

DW (2002): Similar findings for all cities

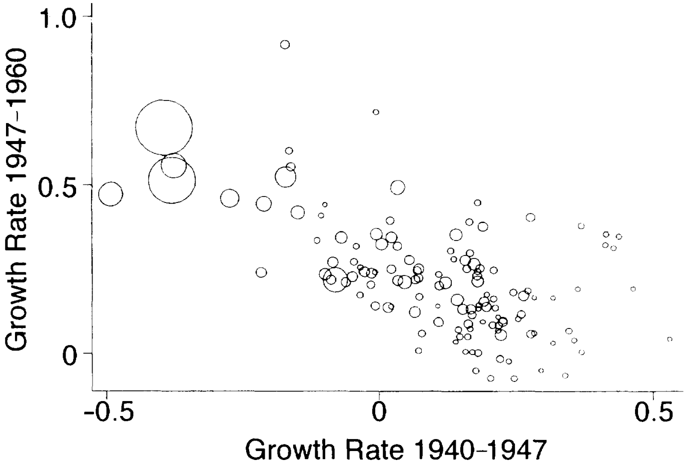


FIGURE 1. EFFECTS OF BOMBING ON CITIES WITH MORE THAN 30,000 INHABITANTS

DW (2008): investigation of city-industry-level component of same shock

TABLE 1: Evolution of Japanese Manufacturing During World War II
(Quantum Indices from Japanese Economic Statistics)

	1941	1946	Change
Manufacturing	206.2	27.4	−87%
Machinery	639.2	38.0	−94%
Metals	270.2	20.5	−92%
Chemicals	252.9	36.9	−85%
Textiles and Apparel	79.4	13.5	−83%
Processed Food	89.9	54.2	−40%
Printing and Publishing	133.5	32.7	−76%
Lumber and Wood	187.0	91.6	−51%
Stone, Clay, Glass	124.6	29.4	−76%

DW (2008): checking whether there is some imperfect correlation within cities

TABLE 2: Correlation of Growth Rates of Industries Within Cities 1938–1948

	Machinery	Metals	Chemicals	Textiles	Food	Printing	Lumber
Metals	0.60						
Chemicals	0.30	0.36					
Textiles	0.12	0.35	0.25				
Food	0.32	0.65	0.31	0.49			
Printing	0.11	0.30	0.04	0.29	0.35		
Lumber	0.23	0.35	0.21	0.25	0.25	0.41	
Ceramics	0.13	0.53	0.36	0.38	0.50	0.41	0.23

DW (2008): not just drop in output, but destruction of assets too

TABLE 3: Inflation Adjusted Percent Decline in Assets Between 1935 and 1945

	Decline
Total	25.4
Buildings	24.6
Harbors and canals	7.5
Bridges	3.5
Industrial machinery and equipment	34.3
Railroads and tramways	7.0
Cars	21.9
Ships	80.6
Electric power generation facilities	10.8
Telecommunication facilities	14.8
Water and sewerage works	16.8

Source: Namakamura, Takafusa.and Masayasu Miyazaki.Shiryo, Taiheiyo Senso Higai Chosa Hokoku (1995), pp. 295–296.

DW (2008): "no path dependence" finding seems to hold, on average, across city-industries

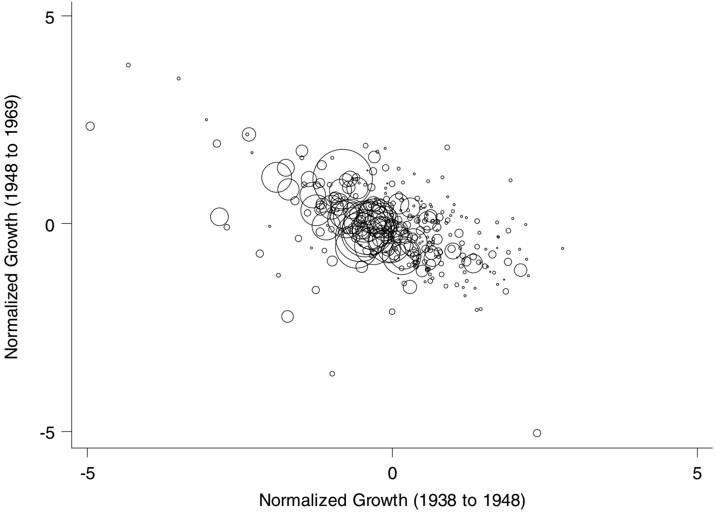


FIGURE 7: Mean-Differenced Industry Growth Rates.

DW (2008):...and even (a little more suggestively) within industries

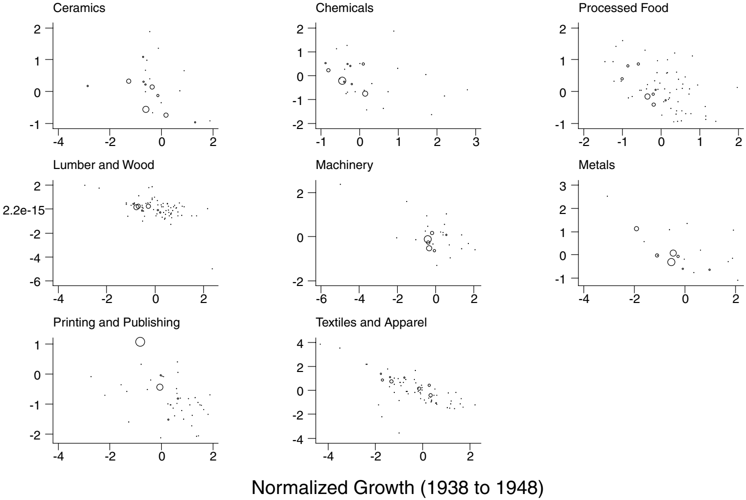


FIGURE 8: Prewar vs Postwar Growth Rate.

Plan for Today's Lecture

- Broad goal: different approaches to path dependence in economic geography settings
- **Papers that look for direct evidence of path dependence:**
  1. WWII bombing: Davis and Weinstein (2002) and Davis and Weinstein (2008)
  2. **Portage: Bleakley and Lin (2012)**
- Quantitative calculations about the importance of path dependence in the US: Allen and Donaldson (2020)

Bleakley and Lin (QJE, 2012)

- BL (2012) pursue a different strategy to look for PD behavior
- Examine a (much) slower withdrawal of fundamentals than in BW (2002/8)—a gradual change in technology that removed the natural "fundamentals" that certain locations enjoyed.
- In particular, BL (2012) look at 'portage sites':
  - Locations in the US where portage (i.e. the trans-shipment of goods from one type of boat to another type of boat) took place before the construction of canals/railroads.
  - Prior to canals/railroads, portage was extremely labor-intensive so portage sites were a source of relatively high labor demand.
- To pin-point exogenous locations of portage sites, BL (2012) use the 'fall line':
  - Geological feature indicating the point at which (in the US) navigable rivers leaving the ocean would first become unnavigable (i.e. at which rapids/waterfalls first emerge)

Bleakley and Lin (2010): The Fall Line

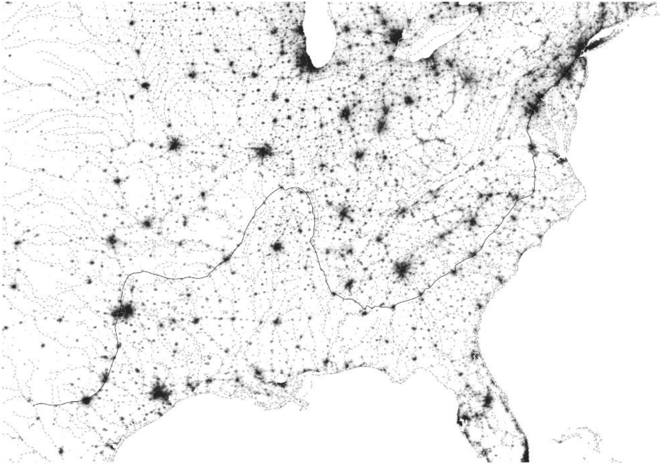


FIGURE A.1
The Density Near Fall-Line/River Intersections
This map shows the contemporary distribution of economic activity across the southeastern United States measured by the 2003 nighttime lights layer. For information on sources, see notes for Figures II and IV.

Bleakley and Lin (2012): The Fall Line

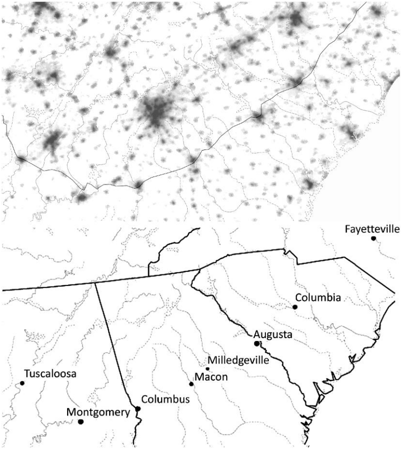


FIGURE II

Fall-Line Cities from Alabama to North Carolina

The map in the upper panel shows the contemporary distribution of economic activity across the southeastern United States, measured by the 2003 nighttime lights layer from NationalAtlas.gov. The nighttime lights are used to present a nearly continuous measure of present-day economic activity at a high spatial frequency. The fall line (solid) is digitized from *Physical Divisions of the United States*, produced by the U.S. Geological Survey. Major rivers (dashed gray) are from NationalAtlas.gov, based on data produced by the United States Geological Survey. Contemporary fall-line cities are labeled in the lower panel.

Bleakley and Lin (2012): The Fall Line

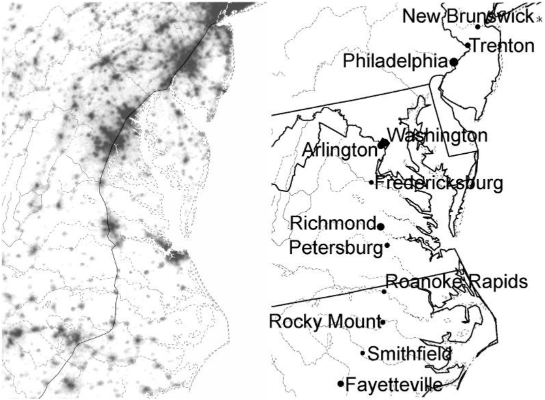


FIGURE IV

Fall-Line Cities from North Carolina to New Jersey

The map in the left panel shows the contemporary distribution of economic activity across the southeastern United States measured by the 2003 nighttime lights layer from NationalAtlas.gov. The nighttime lights are used to present a nearly continuous measure of present-day economic activity at a high spatial frequency. The fall line (solid) is digitized from *Physical Divisions of the United States*, produced by the U.S. Geological Survey. Major rivers (dashed gray) are from NationalAtlas.gov, based on data produced by the U.S. Geological Survey. Contemporary fall-line cities are labeled in the right panel.

Bleakley and Lin (2012): Results

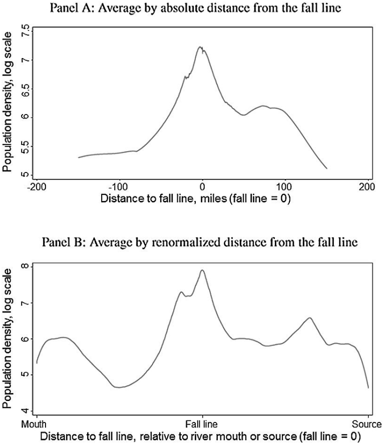


FIGURE III

Population Density in 2000 along Fall-Line Rivers

These graphs display contemporary population density along fall-line rivers. We select census 2000 tracts whose centroids lie within 50 miles along fall-line rivers; the horizontal axis measures distance to the fall line, where the fall line is normalized to zero, and the Atlantic Ocean lies to the left. In Panel A, these distances are calculated in miles. In Panel B, these distances are normalized for each river relative to the river mouth or the river source. The raw population data are then smoothed via Stata's lowess procedure, with bandwidths of 0.3 (Panel A) or 0.1 (Panel B).

Bleakley and Lin (2012): Results

TABLE II

UPSTREAM WATERSHED AND CONTEMPORARY POPULATION DENSITY

	(1) Basic	(2) Other spatial controls	(3)	(4) Water power	(5)
Specifications:		State fixed effects	Distance from various features		
Explanatory variables:					
Panel A: Census Tracts, 2000, N = 21452					
Portage site times upstream watershed	0.467 (0.175)**	0.467 (0.164)***	0.500 (0.114)***	0.496 (0.173)***	0.452 (0.177)**
Binary indicator for portage site	1.096 (0.348)***	1.000 (0.326)***	1.111 (0.219)***	1.099 (0.350)***	1.056 (0.364)***
Portage site times horsepower/100k				−1.812 (1.235)	
Portage site times I(horsepower > 2000)					0.110 (0.311)
Panel B: Nighttime Lights, 1996–97, N = 65000					
Portage site times upstream watershed	0.418 (0.115)***	0.352 (0.102)***	0.456 (0.113)***	0.415 (0.116)***	0.393 (0.111)***
Binary indicator for portage site	0.463 (0.116)***	0.424 (0.111)***	0.421 (0.121)***	0.462 (0.116)***	0.368 (0.132)***
Portage site times horsepower/100k				0.098 (0.433)	
Portage site times I(horsepower > 2000)					0.318 (0.232)
Panel C: Counties, 2000, N = 3480					
Portage site times upstream watershed	0.443 (0.209)**	0.372 (0.185)**	0.423 (0.207)**	0.462 (0.215)**	0.328 (0.154)**
Binary indicator for portage site	0.890 (0.211)***	0.834 (0.194)***	0.742 (0.232)***	0.889 (0.211)***	0.587 (0.210)***
Portage site times horsepower/100k				−0.460 (0.771)	
Portage site times I(horsepower > 2000)					0.991 (0.442)**

Bleakley and Lin (2012): Results

What historical factors are correlated with portage?

TABLE III

PROXIMITY TO HISTORICAL PORTAGE SITE AND HISTORICAL FACTORS

	Baseline	(1) Railroad network length, 1850	(2) Distance to RR hub, 1850	(3) Literate white men, 1850	(4) Literacy rate white men, 1850	(5) College teachers per capita, 1850	(6) Manuf. / agric., 1880	(7) Non-agr. share, 1880	(8) Industrial diversity (1-digit), 1880	(9) Industrial diversity (3-digit), 1880	(10) Water power in use 1885, dummy
Explanatory variables:											
Panel A. Portage and historical factors											
Dummy for proximity		1.451	−0.656	0.557	0.013	0.240	0.065	0.073	0.143	0.927	0.164
to portage site		(0.304)***	(0.254)**	(0.222)**	(0.014)	(0.179)	(0.024)***	(0.025)***	(0.078)*	(0.339)***	(0.053)***
Panel B. Portage and historical factors, conditioned on historical density											
Dummy for proximity		1.023	−0.451	0.021	−0.003	0.213	0.022	0.019	0.033	−0.091	0.169
to portage site		(0.297)***	(0.270)	(0.035)	(0.014)	(0.162)	(0.019)	(0.019)	(0.074)	(0.262)	(0.054)***
Panel C. Portage and contemporary density, conditioned on historical factors											
Dummy for proximity	0.912	0.774	0.751	0.729	0.940	0.883	0.833	0.784	0.847	0.691	0.872
to portage site	(0.236)***	(0.236)***	(0.258)***	(0.187)***	(0.237)***	(0.229)***	(0.227)***	(0.222)***	(0.251)***	(0.221)***	(0.233)***
Historical factor		0.118	−0.098	0.439	0.666	1.349	1.989	2.390	0.838	0.310	0.331
		(0.024)***	(0.022)***	(0.069)***	(0.389)*	(0.164)***	(0.165)***	(0.315)***	(0.055)***	(0.015)***	(0.152)**

Notes. This table displays estimates of equation 1, with the below noted modifications. In Panels A and B, the outcome variables are historical factor densities, as noted in the column headings. The main explanatory variable is a dummy for proximity to a historical portage. Panel B also controls for historical population density. In Panel C, the outcome variable is 2000 population density, measured in natural logarithms, and the explanatory variables are portage proximity and the historical factor density noted in the column heading. Each panel/column presents estimates from a separate regression. The sample consists of all U.S. counties, in each historical year, that are within the watersheds of rivers that cross the fall line. The estimator used is OLS, with standard errors clustered on the 53 watersheds. The basic specification includes a polynomial in latitude and longitude, a set of fixed effects by the watershed of each river that crosses the fall line, and dummies for proximity to the fall line and to a river. Reporting of additional coefficients is suppressed. Data sources and additional variable and sample definitions are found in the text and appendixes.

Bleakley and Lin (2012): Results

Is the portage site effect (today) just the long-lived effect of sunk investments made in the past?

TABLE IV

PROXIMITY TO HISTORICAL PORTAGE SITE AND CONTEMPORARY FACTORS

	(1) Housing units, 1990	(2) Median rents, 1990	(3) Median values, 1990	(4) Interstates, 2000	(5) Major roads, 2000	(6) Rail, 2000	(7) Travel time to work, 1990	(8) Crime, 1995	(9) Born in state, 1990	(10) Water use, 1995	(11) Federal expend., 1997	(12) Gov't. empl., 1997
Explanatory variables:												
Panel A. Portage and contemporary factors												
Dummy for proximity	0.910	0.110	0.108	0.602	0.187	0.858	−0.554	1.224	0.832	0.549	1.063	1.001
to portage site	(0.243)***	(0.040)***	(0.053)**	(0.228)**	(0.071)**	(0.177)***	(0.492)	(0.318)***	(0.186)***	(0.197)***	(0.343)***	(0.283)***
Panel B. Portage and contemporary factors, conditioned on contemporary density												
Dummy for proximity	0.005	0.014	−0.001	0.159	−0.064	0.182	−0.447	−0.007	−0.025	−0.153	0.032	0.114
to portage site	(0.015)	(0.020)	(0.038)	(0.108)	(0.054)	(0.110)	(0.513)	(0.058)	(0.046)	(0.145)	(0.091)	(0.077)

Notes. This table displays estimates of equation (1), with exceptions noted here. In Panels A and B, the outcome variables are current factor densities (natural log of the ratio of quantity per square mile), as noted in the column headings. (The exceptions are house rent and value, which are in logs but not normalized by area, and travel times, which are in minutes.) The coefficient reported is for proximity to historical portage sites. Panel B also controls for current population density. Each cell presents estimates from a separate regression. The sample consists of all US counties, from the indicated year, that are within the watersheds of rivers that cross the fall line. The estimator used is OLS, with standard errors clustered on the 53 watersheds. The specification includes a polynomial in latitude and longitude, a set of fixed effects by the watershed of each river that crosses the fall line, and dummies for proximity to the fall line and to a river. Reporting of additional coefficients is suppressed. Data sources and additional variable and sample definitions are found in the text and appendixes.

Bleakley and Lin (2012): Results

Estimating agglomeration elasticities ("α" in our earlier lectures)

TABLE V

ESTIMATES OF THE EFFECT OF DENSITY ON WAGES USING PORTAGE AS AN INSTRUMENTAL VARIABLE

Log hourly wage	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS
Log population density	0.049	0.085	0.089	0.091
	(0.003)**	(0.032)**	(0.030)**	(0.028)**
Instruments				
Portage-site dummy	–	X	–	X
Log watershed size interaction	–	–	X	X
First-stage statistics				
F	–	8.69	10.7	8.93
p (overidentification)	–	–	–	0.888

Notes. This table displays estimates of regressions of wages on population density. The outcome variable is hourly wage, measured in natural logarithms. Each column presents estimates from a separate regression. The sample consists of all workers in the 2000 IPUMS, age 25–65, that are observed in metropolitan areas in the watersheds of rivers that cross the fall line. In column (1), the estimator used is OLS, with standard errors clustered on the 53 watersheds. In columns (2–4), the estimator used is 2SLS, with standard errors clustered on the 53 watersheds. The basic specification includes, at the worker level, controls for sex, race, ethnicity, nativity, educational attainment, marital status, and age, and, at the area level, a polynomial in latitude and longitude, set of fixed effects for the watershed of each river that crosses the fall line, and dummies for proximity to river and fall line. Two portage-related variables are used as instruments for log population density in this table. The first is a binary indicator for proximity to the river/fall-line intersection. The second is the interaction of portage site with the log of land area in the watershed upstream of the fall line, a variable which proxies for demand for commerce at the portage site. First-stage robust F and p (from a NR^2 Sargan-Hausman overidentification test adjusting for clustering at CONSPUMA level) statistics are also reported in each column. Reporting of additional coefficients is suppressed. Data sources and additional variable and sample definitions are found in the text and appendixes.

Bleakley and Lin (2012): Results

How do historical factors change the portage site effect?

TABLE VI

INTERACTION OF HISTORICAL FACTORS WITH GROWTH AT PORTAGES

	(1) Baseline estimate	(2) Warm climate	(3) College teachers, 1850	(4) Literacy rate, 1850	(5) Industry diversity, 1850	(6) Manuf. / agr., 1880	(7) Regional pop. (donut), 2000
Explanatory variables:							
Dummy for proximity to	0.456	0.727	0.417	0.440	0.346	0.274	0.451
portage site × 20th century	(0.092)***	(0.174)***	(0.092)***	(0.094)***	(0.085)***	(0.085)***	(0.090)***
Additional factor (column		0.124	0.475	−0.731	0.202	0.349	2.843
heading) × 20th century		(0.130)	(0.162)***	(0.218)***	(0.033)***	(0.055)***	(1.626)*
Dummy for portage × add'l		−0.402	1.080	1.083	0.275	0.044	0.034
factor × 20th century		(0.196)**	(0.419)***	(0.472)**	(0.095)***	(0.061)	(0.078)

Notes. This table displays estimates of equation (3) in the text. Each column presents estimates from a separate regression. Each regression uses county-year observations for years 1790–1870 and 1950–2000 and all counties that lie in river watersheds that intersect the fall line. The estimator used is OLS, with standard errors clustered on the 53 watersheds. The outcome variable for each county-year is the natural logarithm of population density, normalized to year 2000 county boundaries. The explanatory variables include a fixed county effect, an indicator variable for the observation year being 1950 or later and its interactions with a spatial trend, a county group indicator, and a portage proximity variable. An additional regressor, noted in column headings, that is interacted with portage proximity and year is also included. These additional variables are transformed to have mean zero with standard deviations displayed in brackets. Reporting of additional coefficients is suppressed. Data sources and additional variable and sample definitions are found in the text and the appendixes.

Lots of "local persistence" studies in Economic History by now

- Surveys: Nunn (2014), Hanlon and Heblich (2020), Lin and Rauch (2020), Voth (2020)
- Examples just within United States:
  - slavery: Nunn (2008)
  - boundaries: Dippel (2014)
  - flooding: Hornbeck and Naidu (2014)
  - mining: Glaeser, Kerr and Kerr (2015)
  - fires: Hornbeck and Keniston (2017)
  - war: Feigenbaum, Lee and Mezzanotti (2018)
  - frontier: Bazzi, Fiszbein and Gebresilasse (2020)
  - immigration: Sequeira, Nunn and Qian (2020)
- Often interpreted as persistence of the shock itself—yet in economic geography models featuring path dependence, *any* feature would persist spatially

Plan for Today's Lecture

- Broad goal: different approaches to path dependence in economic geography settings
- Papers that look for direct evidence of path dependence:
  1. WWII bombing: Davis and Weinstein (2002) and Davis and Weinstein (2008)
  2. Portage: Bleakley and Lin (2012)
- **Quantitative calculations about the importance of path dependence in the US: Allen and Donaldson (2020)**

Allen and Donaldson (2020)

- Develop tools for the quantitative study of path dependence
  - Tractable dynamic model suited to real geography (many regions, unrestricted trade and migration costs)
  - Conditions on parameter values under which model features multiple steady-states, yet equilibrium transition paths unique.
- Estimate parameters using US spatial history (1850-present)
  - Wide range of instruments possible: geography, lagged populations, lagged (and now obsolete) productivity/amenity shifters, responsiveness of economy to temporary shocks
- Attempt to answer counterfactual questions such as:
  - How consequential is path dependence? How bad are chosen steady-states relative to best?

Model setup: Geography

- N locations. Each location $i \in \{1, ..., N\}$ in each time period $t \in \{1, ...\}$ is endowed with:
  - Technology for producing a differentiated good (Armington assumption).
  - An innate productivity $\bar{A}_{it}$.
  - An innate amenity $\bar{u}_{it}$.
- All pairs of locations (i, j) are endowed with:
  - A bilateral trade cost $\tau_{ijt} \geq 1$.
  - A bilateral migration cost $\mu_{ijt} \geq 1$.
- Set $\{\bar{A}_{it}, \bar{u}_{it}, \tau_{ijt}, \mu_{ijt}\}$ comprise the *geography* of the system.

Model setup: Dynamics

- Agents live two periods ("childhood" and "adulthood").

Model setup: Dynamics

- Agents live two periods ("childhood" and "adulthood").
- Consider an agent who is an adult in period t:
  - In period $t-1$, that agent is born where her parent lived.
  - In period t, choose where to live (i.e. produce+consume). Gives birth to child, born in that location.

Model setup: Dynamics

- Agents live two periods ("childhood" and "adulthood").
- Consider an agent who is an adult in period t:
  - In period $t-1$, that agent is born where her parent lived.
  - In period t, choose where to live (i.e. produce+consume). Gives birth to child, born in that location.
- Agents only produce/consume in adulthood, but are *forward-looking* in the dynastic sense – derive utility from their child's consumption.

Model setup: Dynamics

- Agents live two periods ("childhood" and "adulthood").
- Consider an agent who is an adult in period t:
  - In period $t-1$, that agent is born where her parent lived.
  - In period t, choose where to live (i.e. produce+consume). Gives birth to child, born in that location.
- Agents only produce/consume in adulthood, but are *forward-looking* in the dynastic sense – derive utility from their child's consumption.
- Denote by L_{it} the adult population in location i in time t.

Model setup: Dynamics

- Agents live two periods ("childhood" and "adulthood").
- Consider an agent who is an adult in period t:
  - In period $t-1$, that agent is born where her parent lived.
  - In period t, choose where to live (i.e. produce+consume). Gives birth to child, born in that location.
- Agents only produce/consume in adulthood, but are *forward-looking* in the dynastic sense – derive utility from their child's consumption.
- Denote by L_{it} the adult population in location i in time t.
- Initial population distribution $\{L_{i0}\}$ given exogenously.

Model setup: Production

- Locations i make Armington-differentiated (EoS $= \sigma$) goods

Model setup: Production

- Locations i make Armington-differentiated (EoS $= \sigma$) goods

- Ricardian technology, with external economies and trade costs:

$$q_{ijt}(\phi) = \frac{1}{\tau_{ijt}} A_{it} l_{it}(\phi), \qquad A_{it} = \bar{A}_{it} L_{it}^{\alpha_1} L_{it-1}^{\alpha_2}, \qquad L_{it} \equiv \int l_{it}(\phi) d\phi$$

Model setup: Production

- Locations i make Armington-differentiated (EoS $= \sigma$) goods

- Ricardian technology, with external economies and trade costs:

$$q_{ijt}(\phi) = \frac{1}{\tau_{ijt}} A_{it} l_{it}(\phi), \quad A_{it} = \bar{A}_{it} L_{it}^{\alpha_1} L_{it-1}^{\alpha_2}, \quad L_{it} \equiv \int l_{it}(\phi) d\phi$$

- Parameter α_1:
  - Many known microfoundations
    - $\alpha_1 > 0$: e.g. Marshallian externalities; monopolistic competition with free entry
    - $\alpha_1 < 0$: e.g. fixed factors with Cobb-Douglas technology

Model setup: Production

- Locations i make Armington-differentiated (EoS $= \sigma$) goods

- Ricardian technology, with external economies and trade costs:

$$q_{ijt}(\phi) = \frac{1}{\tau_{ijt}} A_{it} l_{it}(\phi), \quad A_{it} = \bar{A}_{it} L_{it}^{\alpha_1} L_{it-1}^{\alpha_2}, \quad L_{it} \equiv \int l_{it}(\phi) d\phi$$

- Parameter α_1:
  - Many known microfoundations
    - $\alpha_1 > 0$: e.g. Marshallian externalities; monopolistic competition with free entry
    - $\alpha_1 < 0$: e.g. fixed factors with Cobb-Douglas technology

- Parameter α_2:
  - One possible microfoundation: durable innovations with dissipation of rents to innovators after one period (Deneckere and Judd, 1992; Desmet and Rossi-Hansberg, 2014)

Model setup: Consumption

- Notation: let $Z \equiv \exp(\widetilde{Z})$ for any variable Z...
- Adult in location i in year t enjoys utility:

$$\widetilde{V}_{it} = \log W_{it} + \delta E_t \left[max_j \left\{ \widetilde{V}_{j,t+1} - \tilde{\mu}_{ij,t+1} + \epsilon_{ij,t+1} \right\} \right], \quad \epsilon_t \sim Frechet(\theta)$$

- Or equivalently:

$$V_{it} = W_{it} \Pi^{\delta}_{i,t+1}$$

  - where $\Pi_{i,t+1} \equiv \left(\sum_{k=1}^{N} \left(V_{k,t+1} / \mu_{ik,t+1} \right)^{\theta} \right)^{\frac{1}{\theta}}$ is expected utility of a child born in location i in year t.

Model setup: Consumption

- Notation: let $Z \equiv \exp(\widetilde{Z})$ for any variable Z...
- Adult in location i in year t enjoys utility:

$$\widetilde{V}_{it} = \log W_{it} + \delta E_t \left[max_j \left\{ \widetilde{V}_{j,t+1} - \tilde{\mu}_{ij,t+1} + \epsilon_{ij,t+1} \right\} \right], \quad \epsilon_t \sim Frechet(\theta)$$

- Or equivalently:

$$V_{it} = W_{it} \Pi^{\delta}_{i,t+1}$$

  - where $\Pi_{i,t+1} \equiv \left(\sum_{k=1}^{N} \left(V_{k,t+1} / \mu_{ik,t+1} \right)^{\theta} \right)^{\frac{1}{\theta}}$ is expected utility of a child born in location i in year t.

- With period utility W_{it} given by:

$$W_{it} = u_{it} \frac{w_{it}}{P_{it}}, \quad u_{it} = \bar{u}_{it} L^{\beta_1}_{it} L^{\beta_2}_{it-1}$$

  - P_{it} is CES price index
  - β_1 and β_2 have analogous microfoundations to production case (spillovers, fixed factors, durable investments with dissipated rents)

Gravity equations for flows

- **Flow of goods (trade):** consumer maximization yields:

$$X_{ijt} = \tau_{ijt}^{1-\sigma} \left(\frac{w_{it}}{A_{it}} \right)^{1-\sigma} P_{jt}^{\sigma-1} w_{jt} L_{jt}.$$

  - With $P_{it} \equiv \left(\sum_{k=1}^{N} \left(\tau_{kit} \frac{w_{kt}}{A_{kt}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$

Gravity equations for flows

- **Flow of goods (trade):** consumer maximization yields:

$$X_{ijt} = \tau_{ijt}^{1-\sigma} \left(\frac{w_{it}}{A_{it}} \right)^{1-\sigma} P_{jt}^{\sigma-1} w_{jt} L_{jt}.$$

  - With $P_{it} \equiv \left(\sum_{k=1}^{N} \left(\tau_{kit} \frac{w_{kt}}{A_{kt}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$

- **Flow of people (migration):** measure of children born in location i in year $t-1$ moving to location j as adults follows:

$$L_{ijt} = \mu_{ijt}^{-\theta} \Pi_{it}^{-\theta} L_{it-1} V_{jt}^{\theta},$$

Equilibrium conditions

For any initial population $\{L_{i0}\}$ and geography $\{\bar{A}_{it}, \bar{u}_{it}, \tau_{ijt}, \mu_{ijt}\}$, an **equilibrium** is $\{L_{it} > 0, w_{it}, W_{it}, \Pi_{it}, V_{it}\}$ s.t. $\forall i, t$:

1. Payments to labor are equal to total sales: $w_{it} L_{it} = \sum_{j=1}^{N} X_{ijt}$
2. Trade is balanced: $w_{it} L_{it} = \sum_{j=1}^{N} X_{jit}$
3. Contemporaneous population is equal to total immigration: $L_{it} = \sum_{j=1}^{N} L_{jit}$
4. Historical population is equal to total emigration: $L_{it-1} = \sum_{j=1}^{N} L_{ijt}$
5. Agents are forward-looking: $V_{it} = W_{it} \Pi_{i.t+1}^{\delta}$

Equilibrium conditions

For any initial population $\{L_{i0}\}$ and geography $\{\bar{A}_{it}, \bar{u}_{it}, \tau_{ijt}, \mu_{ijt}\}$, an **equilibrium** is $\{L_{it} > 0, w_{it}, W_{it}, \Pi_{it}, V_{it}\}$ s.t. $\forall i, t$:

1. Payments to labor are equal to total sales: $w_{it} L_{it} = \sum_{j=1}^{N} X_{ijt}$
2. Trade is balanced: $w_{it} L_{it} = \sum_{j=1}^{N} X_{jit}$
3. Contemporaneous population is equal to total immigration: $L_{it} = \sum_{j=1}^{N} L_{jit}$
4. Historical population is equal to total emigration: $L_{it-1} = \sum_{j=1}^{N} L_{ijt}$
5. Agents are forward-looking: $V_{it} = W_{it} \Pi_{i.t+1}^{\delta}$

A **steady state** for geography $\{\bar{A}_i, \bar{u}_i, \tau_{ij}, \mu_{ij}\}$ is a time-invariant equilibrium $\{L_i > 0, w_i, W_i, \Pi_i, V_i\}$.

Equilibrium conditions + gravity

1. Payments to labor are equal to total sales:

$$w_{it}^{\sigma} L_{it}^{1+\alpha_1(1-\sigma)} = \sum_j \left(\frac{\bar{A}_{it} L_{it-1}^{\alpha_2} \bar{u}_{jt} L_{jt-1}^{\beta_2}}{\tau_{ijt}} \right)^{\sigma-1} W_{jt}^{1-\sigma} w_{jt}^{\sigma} L_{jt}^{1+\beta_1(\sigma-1)}$$

2. Trade is balanced:

$$w_{it}^{1-\sigma} L_{it}^{\beta_1(1-\sigma)} W_{it}^{\sigma-1} = \sum_j \left(\frac{\bar{u}_{it} L_{it-1}^{\beta_2} \bar{A}_{jt} L_{jt-1}^{\alpha_2}}{\tau_{jit}} \right)^{\sigma-1} w_{jt}^{1-\sigma} L_{jt}^{\alpha_1(\sigma-1)}$$

3. Contemporaneous population is equal to total immigration:

$$L_{it} V_{it}^{-\theta} = \sum_j \mu_{jit}^{-\theta} \Pi_{jt}^{-\theta} L_{jt-1},$$

4. Historical population is equal to total emigration:

$$\Pi_{it}^{\theta} \equiv \sum_j \mu_{ijt}^{-\theta} V_{jt}^{\theta}$$

5. Agents are forward-looking: $V_{it} = W_{it} \Pi_{i,t+1}^{\delta}$

A useful new result

- Consider system of $N \times H \times T$ **nonlinear** equations in $x > 0$ (for known $K \geq 0$ and ε):

$$x_{i,h,t} = \sum_{j=1}^{N} K_{ij,h,t} \prod_{h'=1}^{H} \left(x_{j,h',t}^{\varepsilon_{h,h'}^{j,t}} \right) \left(x_{j,h',t+1}^{\varepsilon_{h,h'}^{j,t+1}} \right) \left(x_{i,h',t+1}^{\varepsilon_{h,h'}^{i,t+1}} \right) \left(x_{j,h',t-1}^{\varepsilon_{h,h'}^{j,t-1}} \right) \left(x_{i,h',t-1}^{\varepsilon_{h,h'}^{i,t-1}} \right)$$

A useful new result

- Consider system of $N \times H \times T$ **nonlinear** equations in $x > 0$ (for known $K \geq 0$ and ε):

$$x_{i,h,t} = \sum_{j=1}^{N} K_{ij,h,t} \prod_{h'=1}^{H} \left(x_{j,h',t}^{\varepsilon_{h,h'}^{\mathtt{j,t}}} \right) \left(x_{j,h',t+1}^{\varepsilon_{h,h'}^{\mathtt{j,t+1}}} \right) \left(x_{i,h',t+1}^{\varepsilon_{h,h'}^{\mathtt{i,t+1}}} \right) \left(x_{j,h',t-1}^{\varepsilon_{h,h'}^{\mathtt{j,t-1}}} \right) \left(x_{i,h',t-1}^{\varepsilon_{h,h'}^{\mathtt{i,t-1}}} \right)$$

- **Theorem 1**: *Let* $\mathbf{E}^{\mathtt{j,t}}$*, etc., denote the* $H \times H$ *matrix with elements* $\varepsilon_{h,h'}^{\mathtt{j,t}}$*. Then the above system has at most one bounded solution if the only bounded solution to the* $H \times T$ **linear** *system*:

$$\left(\left| \mathbf{E}^{\mathtt{i,t-1}} \right| + \left| \mathbf{E}^{\mathtt{j,t-1}} \right| \right) \boldsymbol{\mu}_{t-1} - \left(\mathbf{I} - \left| \mathbf{E}^{\mathtt{j,t}} \right| \right) \boldsymbol{\mu}_t + \left(\left| \mathbf{E}^{\mathtt{j,t+1}} \right| + \left| \mathbf{E}^{\mathtt{i,t+1}} \right| \right) \boldsymbol{\mu}_{t+1} = \boldsymbol{b}_t$$

for $\boldsymbol{b}_t \geq \mathbf{0}$ *satisfying* $\boldsymbol{\mu}_t \geq \mathbf{0}$ *is* $\boldsymbol{b}_t = \boldsymbol{\mu}_t = \mathbf{0}$.

A useful new result

- Consider system of $N \times H \times T$ **nonlinear** equations in $x > 0$ (for known $K \geq 0$ and ε):

$$x_{i,h,t} = \sum_{j=1}^{N} K_{ij,h,t} \prod_{h'=1}^{H} \left(x_{j,h',t}^{\varepsilon_{h,h'}^{\mathtt{j,t}}} \right) \left(x_{j,h',t+1}^{\varepsilon_{h,h'}^{\mathtt{j,t+1}}} \right) \left(x_{i,h',t+1}^{\varepsilon_{h,h'}^{\mathtt{i,t+1}}} \right) \left(x_{j,h',t-1}^{\varepsilon_{h,h'}^{\mathtt{j,t-1}}} \right) \left(x_{i,h',t-1}^{\varepsilon_{h,h'}^{\mathtt{i,t-1}}} \right)$$

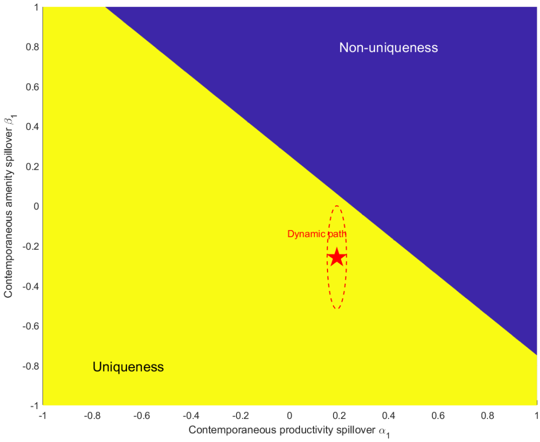
- **Theorem 1**: *Let* $\mathbf{E}^{\mathtt{j,t}}$*, etc., denote the* $H \times H$ *matrix with elements* $\varepsilon_{h,h'}^{\mathtt{j,t}}$*. Then the above system has at most one bounded solution if the only bounded solution to the* $H \times T$ **linear** *system*:

$$\left(\left|\mathbf{E}^{\mathtt{i,t-1}}\right| + \left|\mathbf{E}^{\mathtt{j,t-1}}\right|\right) \boldsymbol{\mu}_{t-1} - \left(\mathbf{I} - \left|\mathbf{E}^{\mathtt{j,t}}\right|\right) \boldsymbol{\mu}_t + \left(\left|\mathbf{E}^{\mathtt{j,t+1}}\right| + \left|\mathbf{E}^{\mathtt{i,t+1}}\right|\right) \boldsymbol{\mu}_{t+1} = \boldsymbol{b}_t$$

for $\boldsymbol{b}_t \geq \mathbf{0}$ *satisfying* $\boldsymbol{\mu}_t \geq \mathbf{0}$ *is* $\boldsymbol{b}_t = \boldsymbol{\mu}_t = \mathbf{0}$*.*

- Extends Allen, Arkolakis and Li (2021) (which require $\rho\left(\left|\mathbf{E}^{\mathtt{j,t}}\right|\right) < 1$) to dynamic systems that arise due to forward-looking agents.

Roughly: uniqueness if *contemporaneous* spillovers (α_1, β_1) are net dispersive



How persistent are historical shocks?

- Define $\chi_{L,t} \equiv \frac{\max_i L_{i,t}/L_{i,t-1}}{\min_i L_{i,t}/L_{i,t-1}}$ for endogenous variable L (similarly $\chi_{V,t}$,$\chi_{\Pi,t}$). $\chi_{L,t}$ is therefore an economy-wide summary of how far L is from steady-state

How persistent are historical shocks?

- Define $\chi_{L,t} \equiv \frac{\max_i L_{i,t}/L_{i,t-1}}{\min_i L_{i,t}/L_{i,t-1}}$ for endogenous variable L (similarly $\chi_{V,t}$,$\chi_{\Pi,t}$). $\chi_{L,t}$ is therefore an economy-wide summary of how far L is from steady-state

- *Persistence:* How are $\chi_t' s$ constrained by $\chi_{t-1}' s$?

How persistent are historical shocks?

- Define $\chi_{L,t} \equiv \frac{\max_i L_{i,t}/L_{i,t-1}}{\min_i L_{i,t}/L_{i,t-1}}$ for endogenous variable L (similarly $\chi_{V,t}$,$\chi_{\Pi,t}$). $\chi_{L,t}$ is therefore an economy-wide summary of how far L is from steady-state

- *Persistence:* How are $\chi_t's$ constrained by $\chi_{t-1}'s$?

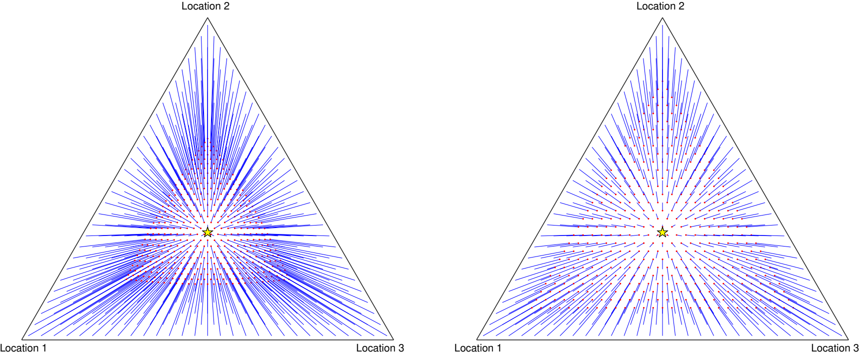
- **Proposition 1**: *Suppose that $\delta = 0$ and $\rho\left(\left|\mathbf{E}^{j,t}\right|\right) < 1$. Then:*

$$\begin{pmatrix} \ln \chi_{L,t} \\ \ln \chi_{V,t} \\ \ln \chi_{\Pi,t} \end{pmatrix} \leq \mathbf{R}(\mathbf{E}^{j,t}) \begin{pmatrix} \ln \chi_{L,t-1} \\ \ln \chi_{V,t-1} \\ \ln \chi_{\Pi,t-1} \end{pmatrix}$$

where $\mathbf{R}(\mathbf{E}^{j,t})$ is a matrix, whose elements depend on $(\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta)$, with the property that $\rho(\mathbf{R}(\mathbf{E}^{j,t})) \to \infty$ as $\rho\left(\left|\mathbf{E}^{j,t}\right|\right) \nearrow 1$.

$N = 3$ example illustrating Proposition 1 (closer to non-uniqueness, more persistence)

- Symmetric locations, $\beta_1 = \beta_2 = \alpha_2 = 0$, but:
  - Left: $\alpha_1 = -0.20$, far from $\rho\left(\left|\mathbf{E}^{j,t}\right|\right) = 1$, $\Rightarrow$ *less persistence*
  - Right: $\alpha_1 = 0$, closer to $\rho\left(\left|\mathbf{E}^{j,t}\right|\right) = 1$, $\Rightarrow$ *more persistence*



When are steady-states unique?

- Define $\alpha_{ss} \equiv \alpha_1 + \alpha_2$ and $\beta_{ss} \equiv \beta_1 + \beta_2$, and matrix:

$$\mathbf{B} \equiv \begin{pmatrix} \left| \frac{1-\frac{\sigma}{\theta}-\beta_{ss}+\alpha_{ss}\sigma+\beta_{ss}\sigma+\frac{1}{\theta}}{\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma} \right| & \left| \frac{(1+\delta)(\alpha_{ss}+1)\left(\frac{\sigma-1}{\theta}\right)}{\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma} \right| \\ \left| \frac{(2\sigma-1)/(\sigma-1)}{\left(\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma\right)} \right| & \left| \frac{1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma-\delta\frac{\sigma}{\theta}}{\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma} \right| \end{pmatrix}$$

When are steady-states unique?

- Define $\alpha_{ss} \equiv \alpha_1 + \alpha_2$ and $\beta_{ss} \equiv \beta_1 + \beta_2$, and matrix:

$$\mathbf{B} \equiv \begin{pmatrix} \left| \frac{1-\frac{\sigma}{\theta}-\beta_{ss}+\alpha_{ss}\sigma+\beta_{ss}\sigma+\frac{1}{\theta}}{\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma} \right| & \left| \frac{(1+\delta)(\alpha_{ss}+1)\left(\frac{\sigma-1}{\theta}\right)}{\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma} \right| \\ \left| \frac{(2\sigma-1)/(\sigma-1)}{\left(\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma\right)} \right| & \left| \frac{1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma-\delta\frac{\sigma}{\theta}}{\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma} \right| \end{pmatrix}$$

- **Proposition 2(a)**: *Under symmetric trade and migration costs, there exists a steady-state equilibrium and that equilibrium is unique if* $\rho(\mathbf{B}) < 1$.

When are steady-states unique?

- Define $\alpha_{ss} \equiv \alpha_1 + \alpha_2$ and $\beta_{ss} \equiv \beta_1 + \beta_2$, and matrix:

$$\mathbf{B} \equiv \begin{pmatrix} \left| \frac{1-\frac{\sigma}{\theta}-\beta_{ss}+\alpha_{ss}\sigma+\beta_{ss}\sigma+\frac{1}{\theta}}{\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma} \right| & \left| \frac{(1+\delta)(\alpha_{ss}+1)\left(\frac{\sigma-1}{\theta}\right)}{\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma} \right| \\ \left| \frac{(2\sigma-1)/(\sigma-1)}{\left(\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma\right)} \right| & \left| \frac{1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma-\delta\frac{\sigma}{\theta}}{\frac{\sigma}{\theta}+1-\alpha_{ss}(\sigma-1)-\beta_{ss}\sigma} \right| \end{pmatrix}$$

- **Proposition 2(a)**: *Under symmetric trade and migration costs, there exists a steady-state equilibrium and that equilibrium is unique if* $\rho(\mathbf{B}) < 1$.

- **Proposition 2(b)**: *If* $\rho(\mathbf{B}) > 1$, *many geographies will have multiple steady-states.*

Well-behaved path dependence

- **Recall from Theorem 1**: low (α_1, β_1) is a force for uniqueness of equilibrium.
- **Recall from Proposition 2**: high $(\alpha_1 + \alpha_2, \beta_1 + \beta_2)$ is a force for multiple steady-states.

Well-behaved path dependence

- **Recall from Theorem 1**: low (α_1, β_1) is a force for uniqueness of equilibrium.
- **Recall from Proposition 2**: high $(\alpha_1 + \alpha_2, \beta_1 + \beta_2)$ is a force for multiple steady-states.
- **Combined**: low (α_1, β_1) and high (α_2, β_2) allows for unique dynamics, but (potentially) multiple steady states $\Rightarrow$ "well-behaved" *path dependence*

Well-behaved path dependence

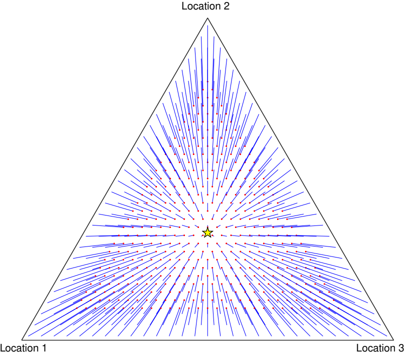
- **Recall from Theorem 1**: low (α_1, β_1) is a force for uniqueness of equilibrium.
- **Recall from Proposition 2**: high $(\alpha_1 + \alpha_2, \beta_1 + \beta_2)$ is a force for multiple steady-states.
- **Combined**: low (α_1, β_1) and high (α_2, β_2) allows for unique dynamics, but (potentially) multiple steady states $\Rightarrow$ "well-behaved" *path dependence*

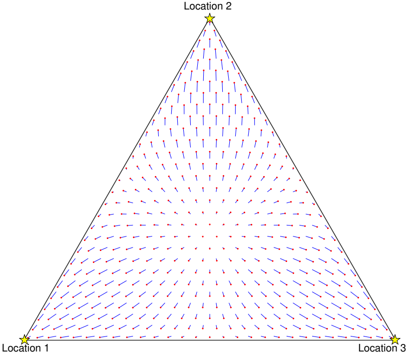
Well-behaved path dependence

- **Recall from Theorem 1**: low (α_1, β_1) is a force for uniqueness of equilibrium.
- **Recall from Proposition 2**: high $(\alpha_1 + \alpha_2, \beta_1 + \beta_2)$ is a force for multiple steady-states.
- **Combined**: low (α_1, β_1) and high (α_2, β_2) allows for unique dynamics, but (potentially) multiple steady states $\Rightarrow$ "well-behaved" *path dependence*
- **Implications:**
  - initial conditions ("history") $\{L_{i0}\}$ determine which SS prevails
  - each SS has its own *basin of attraction*
  - temporary shocks prior to "date 0" can push $\{L_{i0}\}$ across a basin's boundary and hence have permanent consequences

$N = 3$ example illustrating Proposition 2 (historical spillovers can lead to path dependence)

- Symmetric locations, $\beta_1 = \beta_2 = \alpha_1 = 0$, but:
  - Left: $\alpha_2 = 0$, $\Rightarrow \rho(\mathbf{B}) \leq 1$, $\Rightarrow$ *no path dependence*
  - Right: $\alpha_2 = 0.2$, $\Rightarrow \rho(\mathbf{B}) > 1$, $\Rightarrow$ *path dependence*





Data: need L_{ijt}, w_{it} and X_{ijt}

- Time periods t: 1800 (only for $\{L_{i0}\}$), 1850, 1900, 1950 and 2000
- Locations i: 4,975 time-invariant $\cap$ of U.S. counties
- L_{ijt}—population and migration flows:
  - US Census (5% sample microdata) on population by county of current residence and state of birth (and age)
- w_{it}—nominal income per adult:
  - Best available Census proxy for county GDP each each year
- X_{ijt}—bilateral internal trade flows:
  - 1997: Commodity Flow Survey
  - 1949 Crafts-Klein Rail Flow data
  - 1858 and 1900 Chicago Commerce Reports of Trade

Estimation

- Unknown parameters to estimate:
  - 6 key elasticities: $(\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta)$
  - "Level" parameters capturing geographical fundamentals $\left\{\overline{A}_{it}, \overline{u}_{it}, \tau_{ijt}, \mu_{ijt}\right\}$

Estimation

- Unknown parameters to estimate:
  - 6 key elasticities: $(\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta)$
  - "Level" parameters capturing geographical fundamentals $\{\overline{A}_{it}, \overline{u}_{it}, \tau_{ijt}, \mu_{ijt}\}$
- 3-step procedure:

Estimation

- Unknown parameters to estimate:
  - 6 key elasticities: $(\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta)$
  - "Level" parameters capturing geographical fundamentals $\{\overline{A}_{it}, \overline{u}_{it}, \tau_{ijt}, \mu_{ijt}\}$

- 3-step procedure:
  - *Step #1:* Estimate functions connecting trade/migration costs $\{\tau_{ijt}, \mu_{ijt}\}$ to observable components of transportation costs/travel time when using transport network at t.

Estimation

- Unknown parameters to estimate:
  - 6 key elasticities: $(\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta)$
  - "Level" parameters capturing geographical fundamentals $\{\overline{A}_{it}, \overline{u}_{it}, \tau_{ijt}, \mu_{ijt}\}$
- 3-step procedure:
  - *Step #1:* Estimate functions connecting trade/migration costs $\{\tau_{ijt}, \mu_{ijt}\}$ to observable components of transportation costs/travel time when using transport network at t.
  - *Step #2:* recover four key locational characteristics via model inversion

Estimation

- Unknown parameters to estimate:
  - 6 key elasticities: $(\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta)$
  - "Level" parameters capturing geographical fundamentals $\{\overline{A}_{it}, \overline{u}_{it}, \tau_{ijt}, \mu_{ijt}\}$

- 3-step procedure:
  - *Step #1:* Estimate functions connecting trade/migration costs $\{\tau_{ijt}, \mu_{ijt}\}$ to observable components of transportation costs/travel time when using transport network at t.
  - *Step #2:* recover four key locational characteristics via model inversion
  - *Step #3:* linear estimation of labor "supply" and "demand" curves to recover $(\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta)$. Residuals are $\{\overline{A}_{it}, \overline{u}_{it}\}$.

Estimation Step #1: Recovering trade and migration costs

- NLLS procedure :
  - Posit $\ln \tau_{ijt} = \kappa_t^{\tau} \mathit{freight}_{ijt}$ and $\ln \mu_{ijt} = \kappa_t^{\mu} \mathit{time}_{ijt}$
  - Estimate $\mathit{freight}_{ijt}$ and time_{ijt}:
    - Build network of transport possibilities (road, river, rail) in years t
    - Proxy for mode-wise costs in each year using rates/speeds in Fogel (1964), Gordon (2016), Allen and Arkolakis (2014), and Jaworski and Kitchens (2019)
    - Least-"cost" route via fast-marching method
  - NLLS routine to find $\{\kappa_t^{\tau}, \kappa_t^{\mu}\}$ that match best available trade/migration data each year

Estimation Step #1: Recovering trade and migration costs

- NLLS procedure :
  - Posit $\ln \tau_{ijt} = \kappa_t^{\tau} \, freight_{ijt}$ and $\ln \mu_{ijt} = \kappa_t^{\mu} \, time_{ijt}$
  - Estimate $freight_{ijt}$ and $time_{ijt}$:
    - Build network of transport possibilities (road, river, rail) in years t
    - Proxy for mode-wise costs in each year using rates/speeds in Fogel (1964), Gordon (2016), Allen and Arkolakis (2014), and Jaworski and Kitchens (2019)
    - Least-"cost" route via fast-marching method
  - NLLS routine to find $\{\kappa_t^{\tau}, \kappa_t^{\mu}\}$ that match best available trade/migration data each year

	(1)	(2)	(3)	(4)
	1850	1900	1950	2000
	Panel (a): Migration costs			
Effect of travel times	0.019***	0.055***	0.073***	0.066***
on migration flows (κ_t^{μ})	(0.002)	(0.001)	(0.002)	(0.001)
Source	1850 Census	1900 Census	1950 Census	2000 Census
	(1% sample)	(5% sample)	(1% sample)	(5% sample)
Observations	8019	39641	3164	16257
	Panel (b): Trade costs			
Effect of freight costs	0.68	0.52	0.27***	0.36***
on trade flows (κ_t^{τ})	(0.428)	(0.403)	(0.013)	(0.010)
Source	1858 Chicago	1900 Chicago	1949 Crafts-Klein	1997 Commodity
	Commerce Report	Commerce Report	Rail Flow Data	Flow Survey
Observations	10	25	1906	2123

Estimation Step #2: Model inversion

- Define $\hat{T}_{ijt} \equiv \widehat{\tau_{ijt}^{1-\sigma}} = \exp\left(\hat{\kappa_t^{\tau}} \times \textit{freight}_{ijt}\right)$, $\hat{M}_{ijt} \equiv \widehat{\mu_{ijt}^{-\theta}} = \exp\left(\hat{\kappa_t^{\mu}} \times \textit{time}_{ijt}\right)$

Estimation Step #2: Model inversion

- Define $\hat{T}_{ijt} \equiv \widehat{\tau_{ijt}^{1-\sigma}} = \exp\left(\hat{\kappa_t^{\tau}} \times freight_{ijt}\right)$,
  $\hat{M}_{ijt} \equiv \widehat{\mu_{ijt}^{-\theta}} = \exp\left(\hat{\kappa_t^{\mu}} \times time_{ijt}\right)$
- Equilibrium conditions:

$$\mathcal{P}_{it}^{1-\sigma} = \sum_j \hat{T}_{ijt} Y_{jt} P_{jt}^{\sigma-1}, \qquad P_{jt}^{\sigma-1} = \sum_i T_{ijt} Y_{jt} \mathcal{P}_{it}^{\sigma-1}$$

$$\left(\Lambda_{it}^{\theta}\right)^{-1} = \sum_j M_{jit} L_{jt-1} \left(\Pi_{jt}^{\theta}\right)^{-1}, \qquad \Pi_{jt}^{\theta} = \sum_i M_{ijt} L_{jt} \Lambda_{it}^{\theta}$$

Estimation Step #2: Model inversion

- Define $\hat{T}_{ijt} \equiv \widehat{\tau_{ijt}^{1-\sigma}} = \exp\left(\hat{\kappa_t^{\tau}} \times freight_{ijt}\right)$, $\hat{M}_{ijt} \equiv \widehat{\mu_{ijt}^{-\theta}} = \exp\left(\hat{\kappa_t^{\mu}} \times time_{ijt}\right)$
- Equilibrium conditions:

$$\mathcal{P}_{it}^{1-\sigma} = \sum_j \hat{T}_{ijt} Y_{jt} P_{jt}^{\sigma-1}, \qquad P_{jt}^{\sigma-1} = \sum_i T_{ijt} Y_{jt} \mathcal{P}_{it}^{\sigma-1}$$

$$\left(\Lambda_{it}^{\theta}\right)^{-1} = \sum_j M_{jit} L_{jt-1} \left(\Pi_{jt}^{\theta}\right)^{-1}, \qquad \Pi_{jt}^{\theta} = \sum_i M_{ijt} L_{jt} \Lambda_{it}^{\theta}$$

- **Proposition 3**: *Given observed* $\left\{Y_{it}, L_{it}, L_{it-1}, \hat{T}_{ijt}, \hat{M}_{ijt}\right\}$*, there exists unique (to-scale) endogenous characteristics* $\left\{\mathcal{P}_{it}^{\sigma-1}, P_{it}^{\sigma-1}, \Lambda_{it}^{\theta}, \Pi_{it}^{\theta}\right\}$. Maps

Estimation Step #3: The estimating equation

- Log-linear (inverse) labor supply and demand system:

$$\ln w_{it} = \left[\alpha_1 \left(\frac{\sigma - 1}{\sigma}\right) - \frac{1}{\sigma}\right] \ln L_{it} + \alpha_2 \left(\frac{\sigma - 1}{\sigma}\right) \ln L_{it-1} + \frac{1}{1-\sigma} \ln \mathcal{P}_{it}^{1-\sigma} + \frac{\sigma - 1}{\sigma} \ln \bar{A}_{it}$$

$$\ln w_{it} = \left(\frac{1}{\theta} - \beta_1\right) \ln L_{it} + (-\beta_2) \ln L_{it-1} + \frac{1}{\theta} \ln \Lambda_{it}^{\theta} + \left(\frac{1}{1-\sigma}\right) \ln P_{it}^{1-\sigma} - \frac{\delta}{\theta} \ln \Pi_{i,t+1}^{\theta} - \ln \bar{u}_{it}$$

Estimation Step #3: The estimating equation

- Log-linear (inverse) labor supply and demand system:

$$\ln w_{it} = \left[\alpha_1 \left(\frac{\sigma - 1}{\sigma}\right) - \frac{1}{\sigma}\right] \ln L_{it} + \alpha_2 \left(\frac{\sigma - 1}{\sigma}\right) \ln L_{it-1} + \frac{1}{1-\sigma} \ln \mathcal{P}_{it}^{1-\sigma} + \frac{\sigma - 1}{\sigma} \ln \bar{A}_{it}$$

$$\ln w_{it} = \left(\frac{1}{\theta} - \beta_1\right) \ln L_{it} + (-\beta_2) \ln L_{it-1} + \frac{1}{\theta} \ln \Lambda_{it}^{\theta} + \left(\frac{1}{1-\sigma}\right) \ln P_{it}^{1-\sigma} - \frac{\delta}{\theta} \ln \Pi_{i,t+1}^{\theta} - \ln \bar{u}_{it}$$

- Notes:
  - Similar to Rosen-Roback idea we saw (in Glaeser and Gottlieb, 2008).
  - Recovered "market access" parameters $\mathcal{P}_{it}^{1-\sigma}$, Λ_{it}^{θ} affect intercept (not slope).
  - $\bar{A}_{it}$ shifts demand (but not supply); $\bar{u}_{it}$ shifts supply (but not demand).
  - 2SLS can identify elasticities $(\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta, \delta)$ and fundamentals $\{\overline{A}_{it}, \overline{u}_{it}\}$, despite potentially severe issue of multiplicity.
  - In theory, (σ, θ, δ) are identified. In practice, very little power.
    - (σ, θ): use estimates from previous literature
    - δ: match return on wealth (1870-2015) in Jorda et al (2019)

Estimation Step #3: Instruments

- **IV for labor demand curve: Changes in amenities**
  - Following Barreca et al (2016), tech. change like AC and heating have improved amenity value of hot/cold locations
  - Interact average maximum (minimum) temperature in hottest (coldest) month with time trends
  - *Exclusion restriction*: Extreme temperature do not affect productivities (cond'l on controls, fixed effects)

Estimation Step #3: Instruments

- **IV for labor demand curve: Changes in amenities**
  - Following Barreca et al (2016), tech. change like AC and heating have improved amenity value of hot/cold locations
  - Interact average maximum (minimum) temperature in hottest (coldest) month with time trends
  - *Exclusion restriction*: Extreme temperature do not affect productivities (cond'l on controls, fixed effects)
- **IV for labor supply curve: Changes in productivity**
  - Following Bustos et al (2016), tech change in intensive cultivation practices
  - Rise of world demand for soy (Roth, 2018)
  - Interact GAEZ/FAO suitability variables with time trends
  - *Exclusion restriction*: Relative local potential yields do not affect amenities (cond'l on controls, fixed effects)

Estimation Step #3: Instruments

- **IV for labor demand curve: Changes in amenities**
  - Following Barreca et al (2016), tech. change like AC and heating have improved amenity value of hot/cold locations
  - Interact average maximum (minimum) temperature in hottest (coldest) month with time trends
  - *Exclusion restriction*: Extreme temperature do not affect productivities (cond'l on controls, fixed effects)

- **IV for labor supply curve: Changes in productivity**
  - Following Bustos et al (2016), tech change in intensive cultivation practices
  - Rise of world demand for soy (Roth, 2018)
  - Interact GAEZ/FAO suitability variables with time trends
  - *Exclusion restriction*: Relative local potential yields do not affect amenities (cond'l on controls, fixed effects)

- **Controls included:**
  - location fixed effects, region-year fixed effects
  - instruments used in the "other" equation

Estimation results: Productivity Spillovers (α)

Table: Estimated productivity spillovers

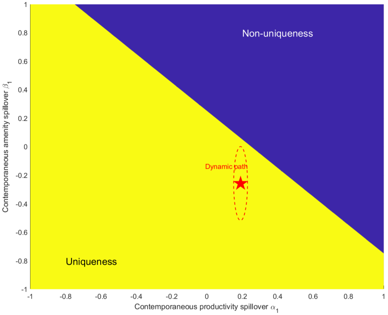
	(1)	(2)	(3)	(4)	(5)	(6)
Estimated parameters:						
Productivity spillover α_1	0.305***	0.334***	0.183***	0.190***	0.143***	0.142***
	(0.039)	(0.045)	(0.035)	(0.040)	(0.034)	(0.039)
Productivity spillover α_2	-0.002	-0.091*	-0.002	-0.041	-0.002	-0.025
	(0.007)	(0.051)	(0.006)	(0.045)	(0.006)	(0.043)
Estimator	OLS	IV	OLS	IV	OLS	IV
Assumed parameters:						
Elasticity of substitution σ	5	5	9	9	13	13
Fixed effects:						
Sub-county	Yes	Yes	Yes	Yes	Yes	Yes
Region-year	Yes	Yes	Yes	Yes	Yes	Yes
Min SW 1st-stage F-stat		58.084		58.084		58.084
Observations	15764	15764	15764	15764	15764	15764

Estimation results: Amenity Spillovers (β)

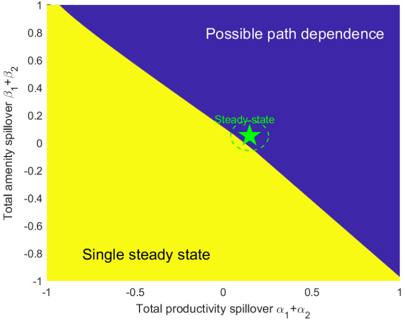
Table: ESTIMATED AMENITY SPILLOVERS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimated parameters:									
Amenity spillover β_1	-0.129	-0.326	-0.392	-0.061	-0.258	-0.324	-0.039	-0.236	-0.301
	(0.266)	(0.264)	(0.263)	(0.267)	(0.265)	(0.265)	(0.267)	(0.266)	(0.266)
Amenity spillover β_2	0.320*	0.314*	0.313*	0.315*	0.310*	0.308*	0.314*	0.309*	0.307*
	(0.180)	(0.178)	(0.178)	(0.180)	(0.178)	(0.178)	(0.180)	(0.179)	(0.179)
Estimator	IV	IV	IV	IV	IV	IV	IV	IV	IV
Assumed parameters:									
Elasticity of substitution σ	5	5	5	9	9	9	13	13	13
Migration elasticity θ	2	4	6	2	4	6	2	4	6
Fixed effects:									
Sub-county	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Min SW 1st-stage F-stat	11.206	11.206	11.206	11.206	11.206	11.206	11.206	11.206	11.206
Observations	15764	15764	15764	15764	15764	15764	15764	15764	15764

Estimates consistent with well-behaved path dependence, substantial persistence



(a) Uniqueness of the transition path



(b) Possibility of path dependence

Counterfactual "perturbation" histories

- Vagaries of spatial location of technological change in "second industrial revolution" (c. 1900)
  - e.g. Detroit became "Motor City" (in part) because Henry Ford born on nearby farm?
  - e.g. Buffalo became "City of Light" (in part) because Edison demonstrates new AC power by illuminating 1901 pan-American Exhibition?

Counterfactual "perturbation" histories

- Vagaries of spatial location of technological change in "second industrial revolution" (c. 1900)
  - e.g. Detroit became "Motor City" (in part) because Henry Ford born on nearby farm?
  - e.g. Buffalo became "City of Light" (in part) because Edison demonstrates new AC power by illuminating 1901 pan-American Exhibition?
- We match locations based on closest in 1900 pop. rank
  - E.g. Buffalo (1900 pop.: 350,000) matches with Cincinnati (1900 pop: 330,000)
  - Milwaukee-New Orleans, Worcester-San Francisco...

Counterfactual "perturbation" histories

- Vagaries of spatial location of technological change in "second industrial revolution" (c. 1900)
  - e.g. Detroit became "Motor City" (in part) because Henry Ford born on nearby farm?
  - e.g. Buffalo became "City of Light" (in part) because Edison demonstrates new AC power by illuminating 1901 pan-American Exhibition?
- We match locations based on closest in 1900 pop. rank
  - E.g. Buffalo (1900 pop.: 350,000) matches with Cincinnati (1900 pop: 330,000)
  - Milwaukee-New Orleans, Worcester-San Francisco...
- For each of 200 counterfactual histories, randomize (without replacement) $\overline{A}_{i,1900}$ between matched location pairs
  - E.g. $\bar{A}_{Mil,1900} = 0.42$, $\bar{A}_{NO,1900} = 0.20$, so $\sim 50\%$ of time, Milwaukee gets unlucky N.O. productivity (and vice versa).

Counterfactual "perturbation" histories

- Vagaries of spatial location of technological change in "second industrial revolution" (c. 1900)
  - e.g. Detroit became "Motor City" (in part) because Henry Ford born on nearby farm?
  - e.g. Buffalo became "City of Light" (in part) because Edison demonstrates new AC power by illuminating 1901 pan-American Exhibition?
- We match locations based on closest in 1900 pop. rank
  - E.g. Buffalo (1900 pop.: 350,000) matches with Cincinnati (1900 pop: 330,000)
  - Milwaukee-New Orleans, Worcester-San Francisco...
- For each of 200 counterfactual histories, randomize (without replacement) $\overline{A}_{i,1900}$ between matched location pairs
  - E.g. $\bar{A}_{Mil,1900} = 0.42$, $\bar{A}_{NO,1900} = 0.20$, so $\sim 50\%$ of time, Milwaukee gets unlucky N.O. productivity (and vice versa).
- Hold constant $\bar{A}_{it}$ for all other years and $(\bar{u}_{it}, \tau_{ijt}, \mu_{ijt})$ for all years.
  - Implication: Differences in economic activity in year 2000 driven only through history (not geography).

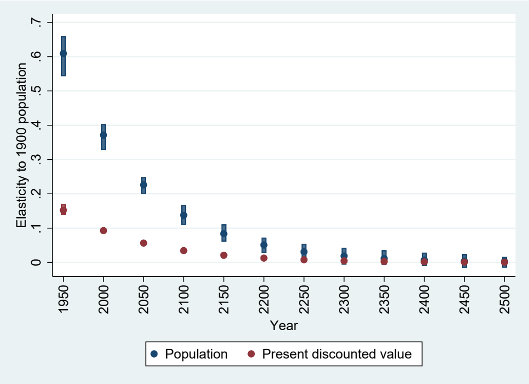
Local persistence elasticities

- Estimate regressions (using "data" for outcome "O" on simulation b) of form
$$\ln O_{it}^{(b)} = \delta_{it}^{O} + \eta_{it}^{O} \ln L_{i,1900}^{(b)} + \varepsilon_{it}^{O(b)}$$
where IV for $L_{i,1900}^{(b)}$ with randomly-drawn $\overline{A}_{i,1900}^{(b)}$

- Call η_{it}^{O} the *local persistence elasticity* for outcome O—estimated causal effect (in this model) of being relatively big in 1900 on outcome at time $t > 1900$

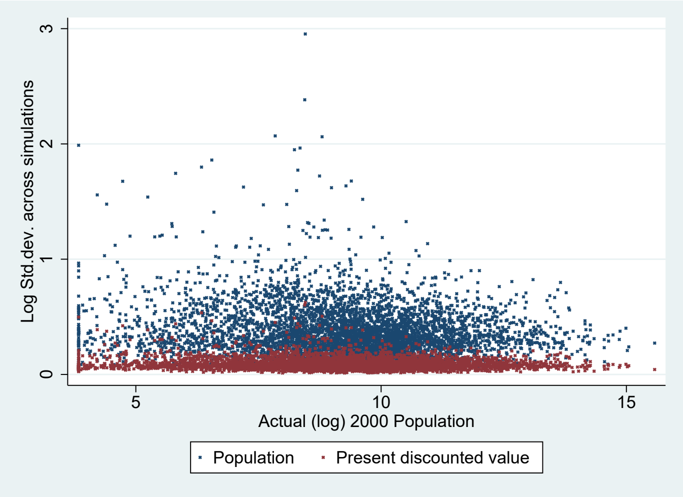
Local persistence elasticities

Figure: How persistent are historical shocks?



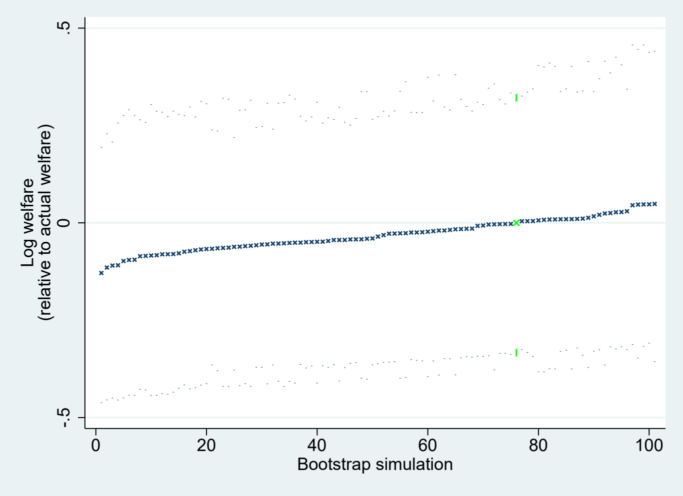
Fragility and resilience

Figure: How resilient are locations to historical shocks?



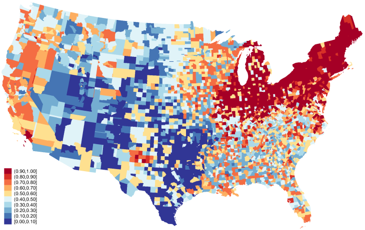
Changes in spatial configuration across simulations

Figure: How lucky was our particular history?

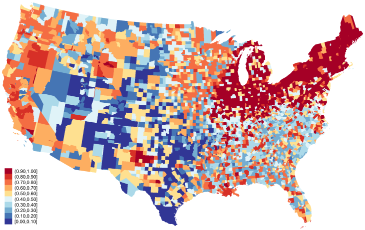


Changes in spatial configuration across simulations

Figure: Which locations were relatively fortunate in their actual history? (Share of simulations with worse outcome than actual.)

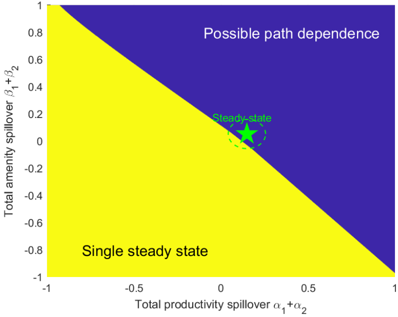


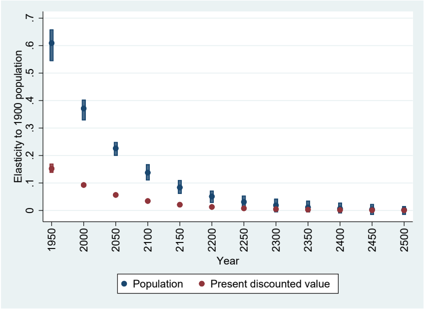
(a) Population



(b) Present discounted value

Can we expect path dependence from these “historical perturbation” shocks?





Can we expect path dependence from these "historical perturbation" shocks?

Aggregate welfare in 3000 across 95% CI of α_2 and β_2 (upper half):