14.582: International Trade II
— Lecture 19 : Economic Geography (Empirics IV)

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 (2018)

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 (2018)

Path Dependence and Economic Geography

- As we've seen so far, evidence for agglomeration economies seems strong:
  - Case studies (e.g. Silicon Valley)
  - Direct estimates (e.g. Berlin Wall, Million Dollar Plants, TVA...)
- Long theoretical tradition highlights implications:
  - Potential for multiple equilibria in static models
  - Potential for *path dependence* (i.e. initial conditions, or long-redundant shocks, still matter for outcomes today) in dynamic models
  - Potential for policies to promote movement to better steady-state
- But is path dependence actually empirically consequential?
  - Should we expect path dependence 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 whether regions/cities' population levels respond to one-off shocks
- The application is to WWII bombing in Japan
- Their findings are surprising and 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 in ways that people are not.

Davis and Weinstein (2002)

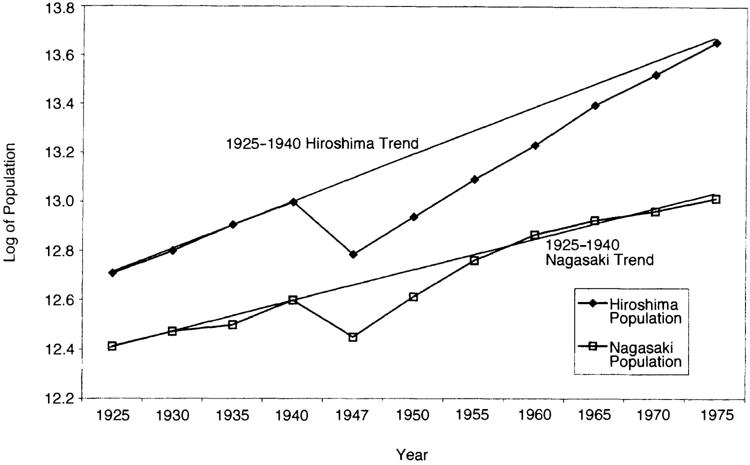


FIGURE 2. POPULATION GROWTH

Davis and Weinstein (2002)

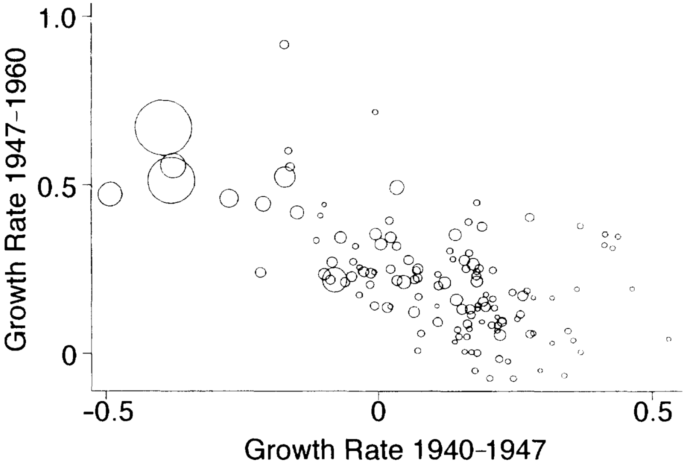


FIGURE 1. EFFECTS OF BOMBING ON CITIES WITH MORE THAN 30,000 INHABITANTS

Davis and Weinstein (2008)

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%

Davis and Weinstein (2008)

Table 2

Correlation of Growth Rates of Industries Within Cities 1938 to 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

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 equipmer	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.295-96.

Davis and Weinstein (2008)

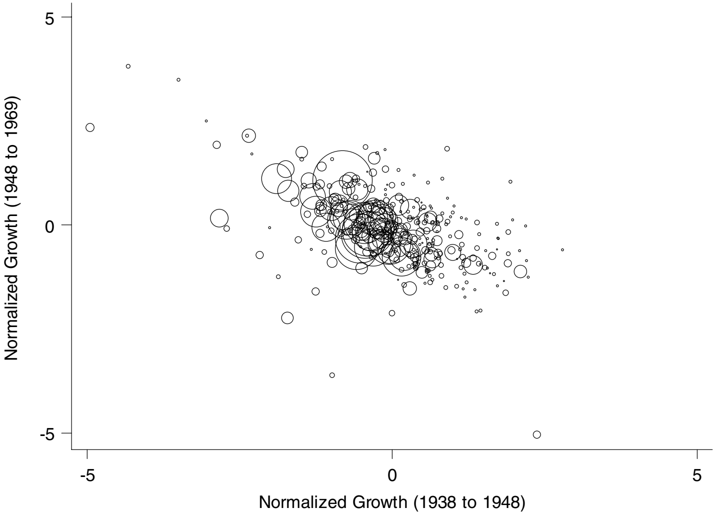


FIGURE 7: Mean-Differenced Industry Growth Rates.

Davis and Weinstein (2008)

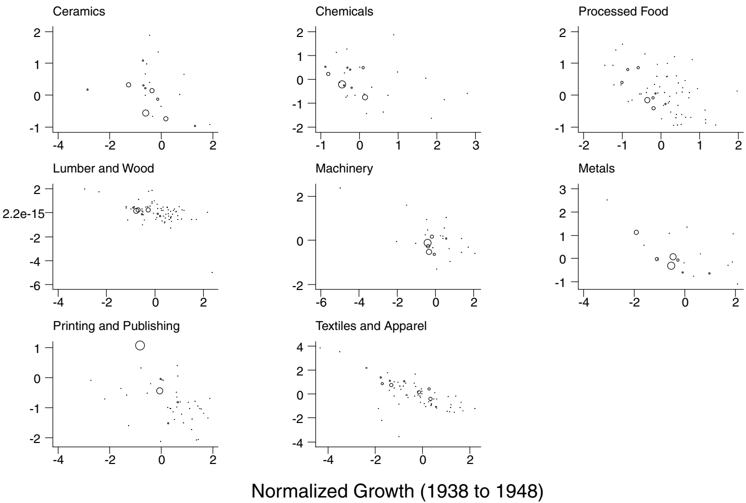


FIGURE 8: Prewar vs Postwar Growth Rate.

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 (2018)

Bleakley and Lin (QJE, 2012)

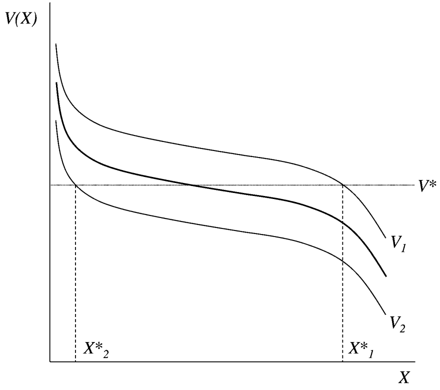
- BL (2012) look for an event that removed a location's natural (i.e. exogenous) productivity advantage/amenity.
- If there are no agglomeration externalities then this location will suffer from this removal.
- But if there are agglomeration externalities then this location might not suffer much at all. Its future success is assured through the logic of multiple equilibria. (This is typically referred to as 'path dependence'.)

Bleakley and Lin (QJE, 2012): Portage

- What is the natural advantage that got removed from some locations?
- BL (2012) look at 'portage sites': locations 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 excess labor demand.
- What is an exogenous source for a portage site? BL (2012) use the 'fall line', a geological feature indicating the point at which (in the US) navigable rivers leaving the ocean would first become unnavigable

Bleakley and Lin (2012): Theory

Panel A: Differences in density with natural advantages and strong congestion costs



Panel B: Differences in density with strong increasing returns

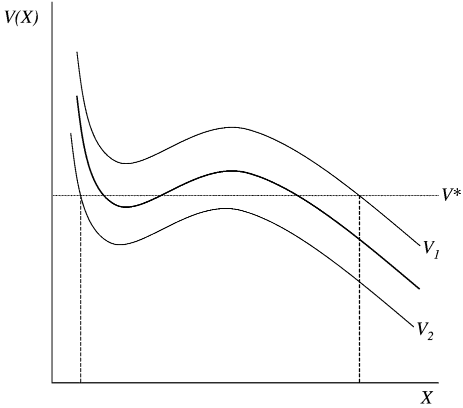


FIGURE VII

Equilibrium Density in a Model with Natural Advantages and Increasing Returns

Bleakley and Lin (2010): The Fall Line

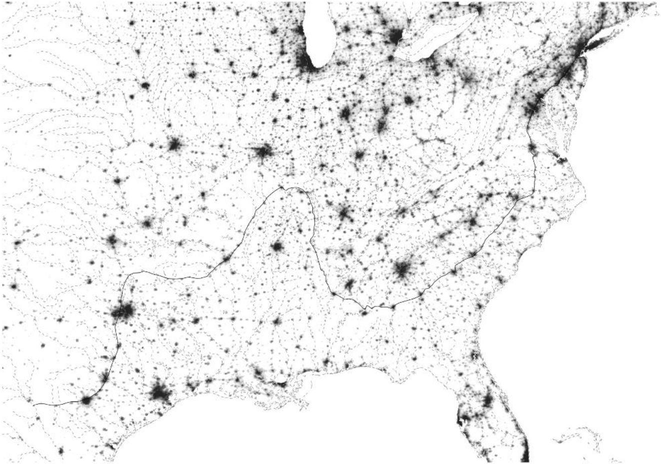


FIGURE A.1

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

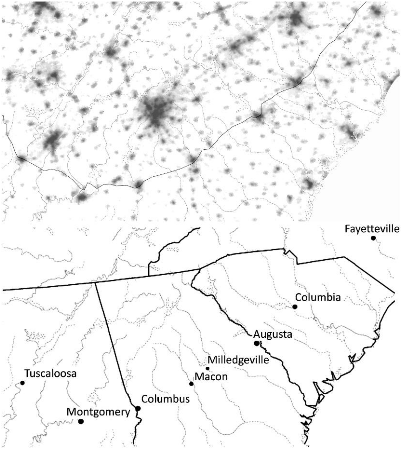


FIGURE II

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

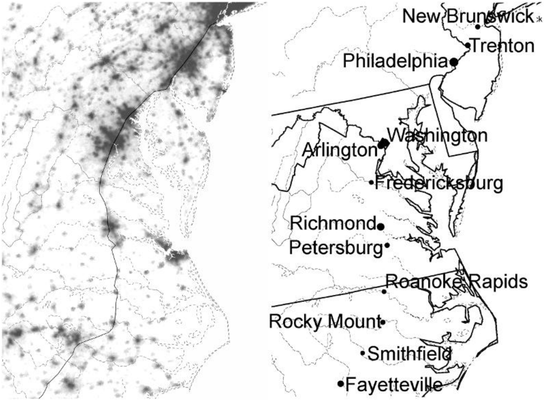


FIGURE IV

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

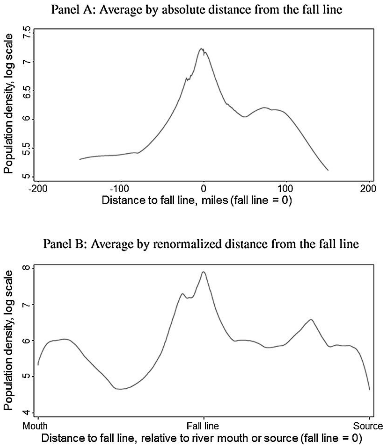


FIGURE III

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 effects

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.

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 (2018)**

Allen and Donaldson (2018)

- 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 (1800-present)
  - Wide range of instruments possible: geography, lagged populations, lagged (and now obsolete) productivity/amenity shifters, responsiveness of economy to temporary shocks
  - For now, preliminary estimates based on one strategy
- Answer counterfactual questions such as:
  - How consequential is path dependence? How bad are chosen steady-states relative to best?
  - Preliminary results suggest path dependence is consequential for the location of economic activity, but not (much) for welfare.

Model overview: Main ingredients

- Flexible bilateral migration and trade frictions, local characteristics.
  - Incorporate real world geography.
- Armington trade, extreme value discrete choice migration.
  - Convenient gravity equations for trade flows and migration.
- Overlapping generations.
  - Straightforward characterization of dynamics.
- Productivity and amenity spillovers.
  - Possibility of multiple steady states and path dependence.

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 iceberg trade cost $\tau_{ijt} \geq 1$.
  - A bilateral iceberg migration cost $\mu_{ijt} \geq 1$.

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 generation $t+1$ in that location.
- Agents only produce/consume in adulthood, do not care about children.
- Let L_{it} be adult population in location i in time t.
- The world's initial population $\{L_{i0}\}$ is given exogenously.

Model setup: Production and Consumption

- Production
  - Perfect competition, (adult) labor only factor of production. Quantity produced:

$$Q_{it} = \underbrace{\left(\bar{A}_{it} L_{it}^{\alpha_1} L_{it-1}^{\alpha_2}\right)}_{\equiv A_{it}} L_{it},$$

    where α_1 and α_2 govern the strength of *contemporaneous* and *historical* productivity spillovers.

- Consumption
  - Adults have CES preferences over differentiated varieties with EoS σ, earn wage w_{it}, have price index P_{it}. Welfare:

$$W_{it} = \underbrace{\left(\bar{u}_{it} L_{it}^{\beta_1} L_{it-1}^{\beta_2}\right)}_{\equiv u_{it}} \frac{w_{it}}{P_{it}},$$

    where β_1 and β_2 govern the strength of *contemporaneous* and *historical* of amenity spillovers.

Gravity

- Armington + consumer maximization yields gravity equation for trade:

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

  - $P_{it} \equiv \left(\sum_{k=1}^{N} \left(\tau_{ki} \frac{w_{kt}}{A_{kt}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$ is CES price index.

- Frechet + welfare maximization yield gravity equation for migration:

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

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

Equilibrium conditions

For any initial population $\{L_{i0}\}$ and geography $\{\bar{A}_{it}, \bar{u}_{it}, \tau_{ijt}, \mu_{ijt}\}$, an equilibrium is $\{L_{it}, w_{it}, W_{it}, \Pi_{it}\}$ 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: $L_{it-1} = \sum_{j=1}^{N} L_{ijt}$

Equilibrium conditions + gravity

Yields $4 \times N \times T$ equations for $4 \times N \times T$ unknowns:

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. The population is equal to total immigration:

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

4. The population is equal to total emigration:

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

Existence and Uniqueness of an Equilibrium

- Define matrix $\mathbf{A}(\alpha_1, \beta_1) \equiv$
$$\begin{pmatrix} \left|\frac{\theta(\alpha_1\sigma+\beta_1(\sigma-1)+1)-(\sigma-1)}{\sigma+\theta(1-(\sigma-1)\alpha_1-\beta_1)}\right| & \left|\frac{\tilde{\sigma}((\sigma-1)(1-(\sigma-1)\alpha_1-\beta_1)+\sigma(\alpha_1\sigma+\beta_1(\sigma-1)+1))}{\sigma+\theta(1-(\sigma-1)\alpha_1-\beta_1)}\right| \\ \left|\frac{\theta}{\tilde{\sigma}(\sigma+\theta(1-(\sigma-1)\alpha_1-\beta_1))}\right| & \left|\frac{\theta(1-(\sigma-1)\alpha_1-\beta_1)}{\sigma+\theta(1-(\sigma-1)\alpha_1-\beta_1)}\right| \end{pmatrix}$$

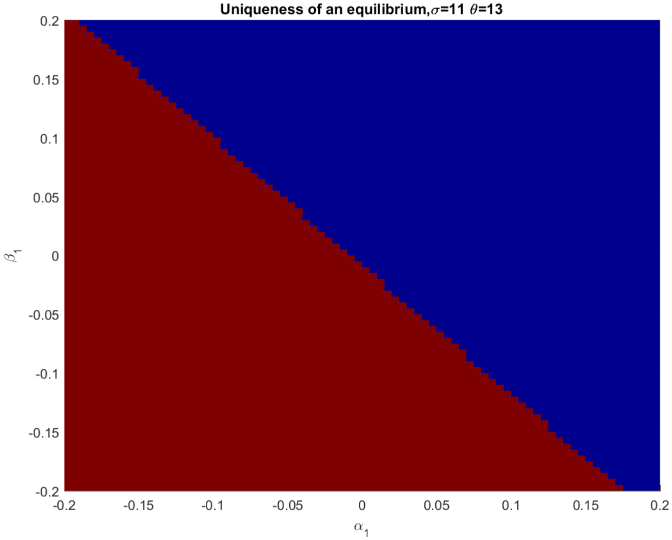
- **Proposition 1(a)**: *For any initial population* $\{L_{i0}\}$ *and geography* $\{\bar{A}_{it} > 0, \bar{u}_{it} > 0, \tau_{ijt} = \tau_{jit}, \mu_{ijt}\}$, *there exists a unique equilibrium if* $\rho(\mathbf{A}(\alpha_1, \beta_1)) \leq 1$.

- This will occur as long as α_1 and β_1 are sufficiently small.

- Proof: System is a special case of the "generalized gravity system" considered in Allen, Arkolakis, Li (2017)

- Note: Result does not depend on values of α_2 and β_2 (since current generation takes L_{it-1} as given).

Existence and Uniqueness of an Equilibrium



Steady state equilibrium

- Consider a steady state with time-invariant geography $\{\bar{A}_i, \bar{u}_i, \tau_{ij}, \mu_{ij}\}$ and endogenous variables $\{L_i, w_i, W_i, \Pi_i\}$.

- **Proposition 1(b)**: *For any geography* $\{\bar{A}_i > 0, \bar{u}_i > 0, \tau_{ij} = \tau_{ji}, \mu_{ij} = \mu_{ji}\}$*, there exists a unique equilibrium if* $\rho(\mathbf{A}(\alpha_1 + \alpha_2, \beta_1 + \beta_2)) \leq 1$*.*

- *Implication*: if $\alpha_1 + \alpha_2 > \alpha_1$ and/or $\beta_1 + \beta_2 > \beta_1$, can have unique transition path but multiple steady states.

Properties of the steady state

- Define steady state welfare as: $\Omega = E\left[\max_i \left(W_i \Pi_i \varepsilon_i\right)\right]$.

- In steady state, welfare is equalized across all locations:

$$W_i \Pi_i L_i^{-\frac{1}{\theta}} = \Omega \ \forall i \in \{1, ..., N\}$$

- Equilibrium steady state distribution of population can be written as:

$$\gamma \ln L_i = C + (1 - \tilde{\sigma}) \ln \bar{u}_i + \tilde{\sigma} \ln \bar{A}_i + (1 - \tilde{\sigma}) \ln \Pi_i - \ln P_i,$$

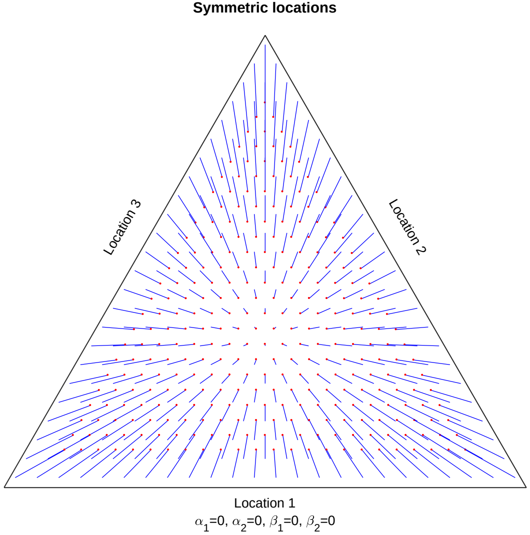
where $\gamma \equiv \frac{1}{\theta}(1 - \tilde{\sigma}) - \frac{\tilde{\sigma}}{\sigma - 1} - (\tilde{\sigma} + 1)\tilde{\beta} + \tilde{\sigma}\tilde{\alpha}$.

- Implication: More people will live in high $\bar{A}_i$, high $\bar{u}_i$, high Π_i, and low P_i places, with elasticities governed by strength of spillovers.

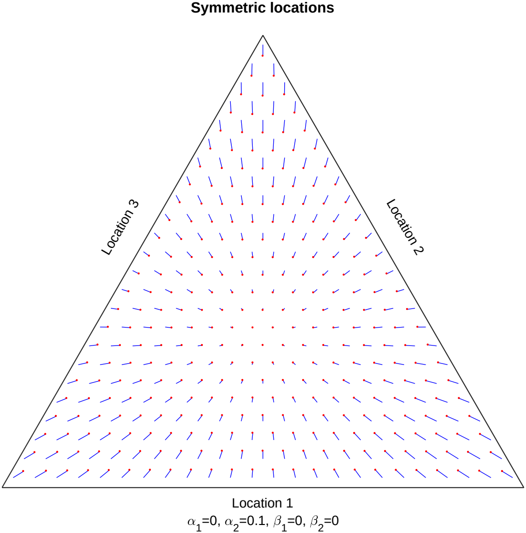
Path dependence

- Suppose $\rho\left(\mathbf{A}\left(\alpha_1, \beta_1\right)\right) \leq 1$ but $\rho\left(\mathbf{A}\left(\alpha_1 + \alpha_2, \beta_1 + \beta_2\right)\right) > 1$.
- Then initial distribution of labor $\{L_{i0}\}$ will determine which steady state the economy converges toward.
- Consider a simple example: 3 identical locations separated by trade costs, with $\alpha_1 = \beta_1 = \beta_2 = 0$, but with increasingly large values of α_2...

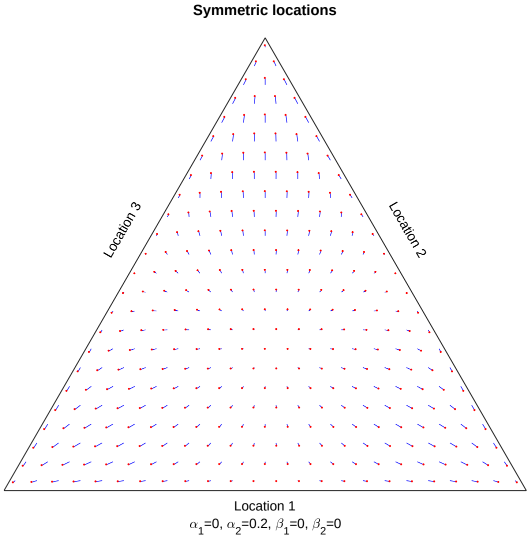
Phase diagram: 3 symmetric locations



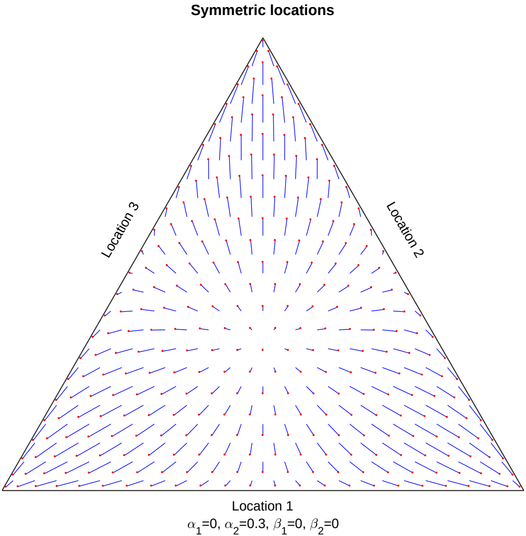
Phase diagram: 3 symmetric locations



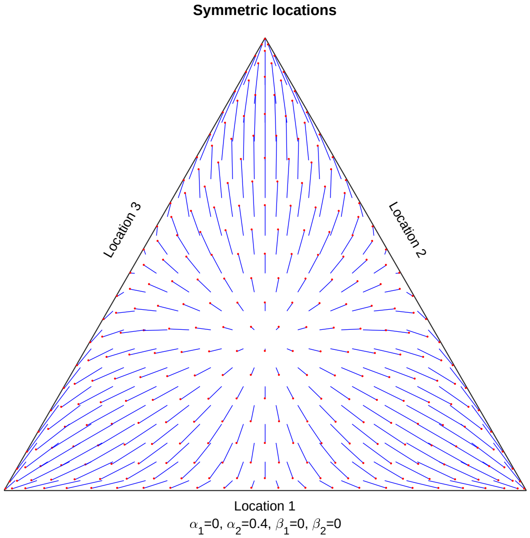
Phase diagram: 3 symmetric locations



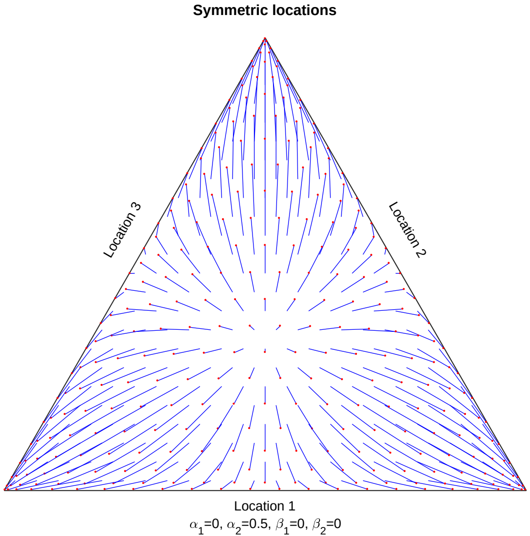
Phase diagram: 3 symmetric locations



Phase diagram: 3 symmetric locations



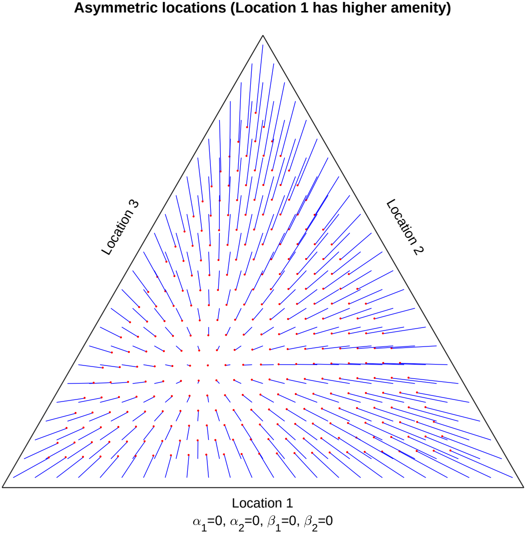
Phase diagram: 3 symmetric locations



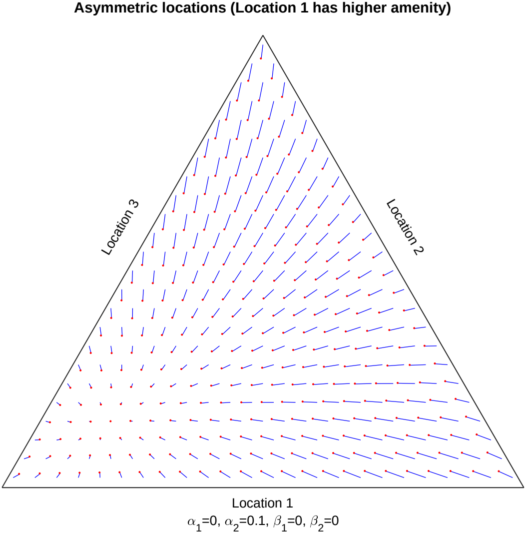
Path dependence: heterogeneous steady states

- In previous example, the 3 stable steady states had identical welfare implications.
- But similar intuition holds when the steady states are associated with different welfare levels.
- Extend previous example to 3 *asymmetric* locations...

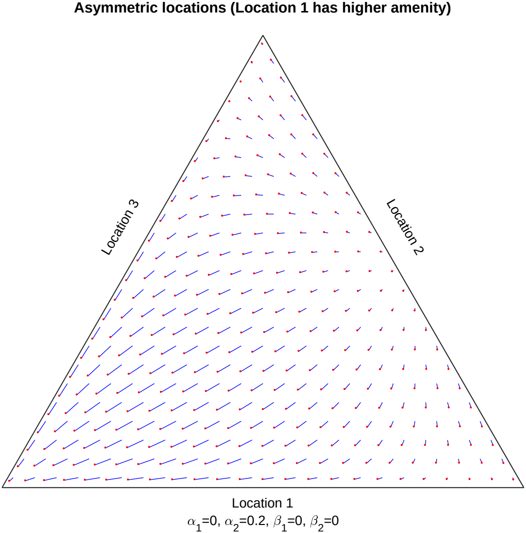
Phase diagram: 3 asymmetric locations



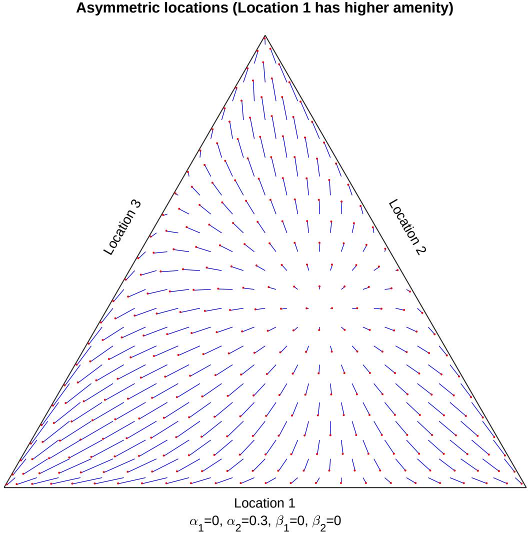
Phase diagram: 3 asymmetric locations



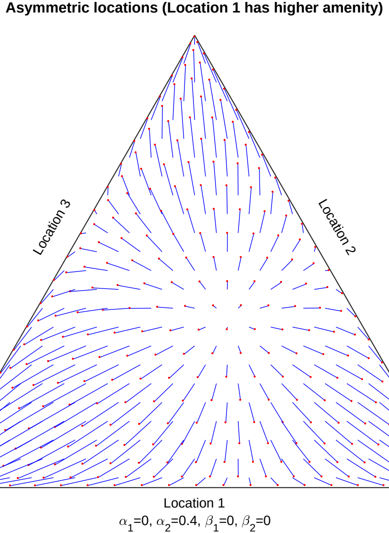
Phase diagram: 3 asymmetric locations



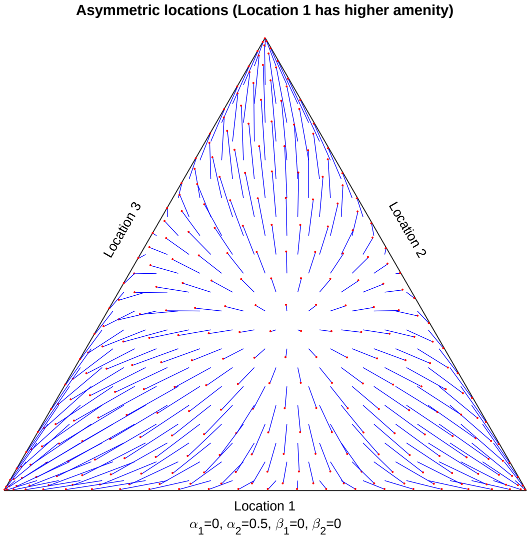
Phase diagram: 3 asymmetric locations



Phase diagram: 3 asymmetric locations



Phase diagram: 3 asymmetric locations



Path dependence: heterogeneous steady states

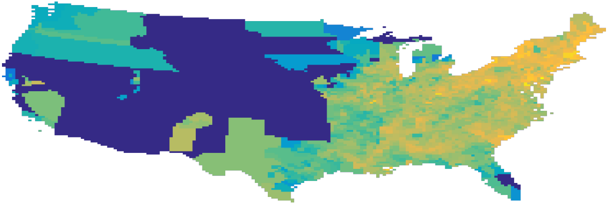
- In previous example, the 3 stable steady states had identical welfare implications.
- But similar intuition holds when the steady states are associated with different welfare levels.
- Extend previous example to 3 *asymmetric* locations.
- *Implication*: Initial population could cause world to converge to "bad" steady state...
- ...but the "good" steady state has larger basin of attraction.

Data

- L_{ijt}:
  - Decennial US Census from 1790-present: data (from 5% sample) on population by county of current residence and state of birth (and age)
  - Manipulate this (with assumptions...) to get proxy for L_{ijt} (and hence $L_{jt} \equiv \sum_i L_{ijt}$)
  - $\{L_{i0}\}$ taken to be populations in 1800
  - Generation will always be 50 years long
- w_{it} :
  - Decennial US Census from 1850-present: data on total agricultural and manufacturing output
  - That plus Cobb-Douglas production function identifies $Y_{it} = w_{it} L_{it}$
- X_{ijt}: From 1997 Commodity Flow Survey.

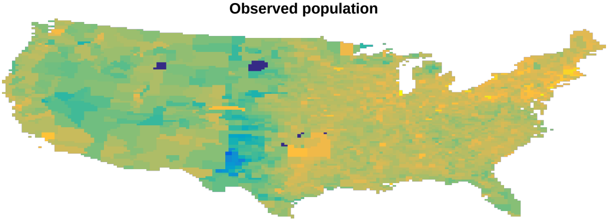
Data on w_{it} and L_{it}: 1850

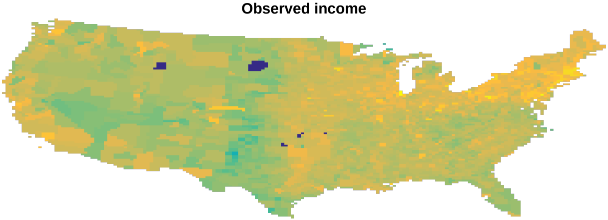
Observed population



Observed income

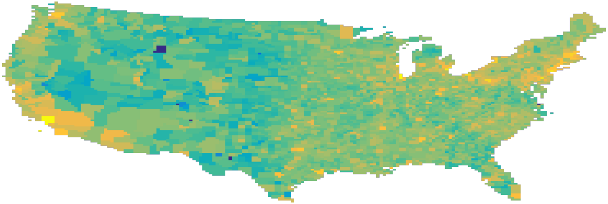
Data on w_{it} and L_{it}: 1900



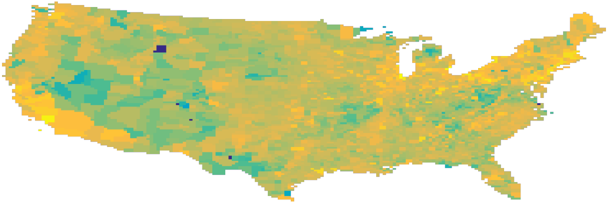


Data on w_{it} and L_{it}: 1950

Observed population

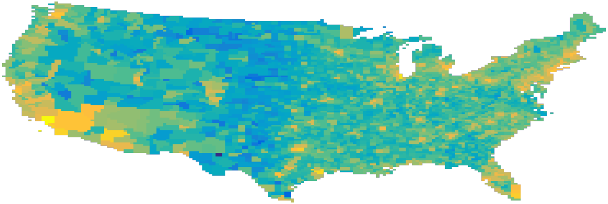


Observed income

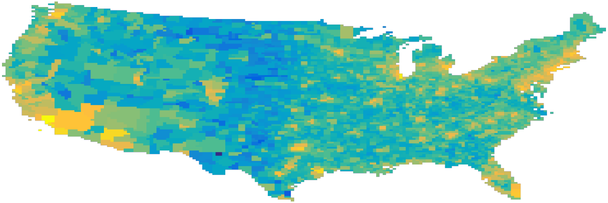


Data on w_{it} and L_{it}: 2000

Observed population



Observed income



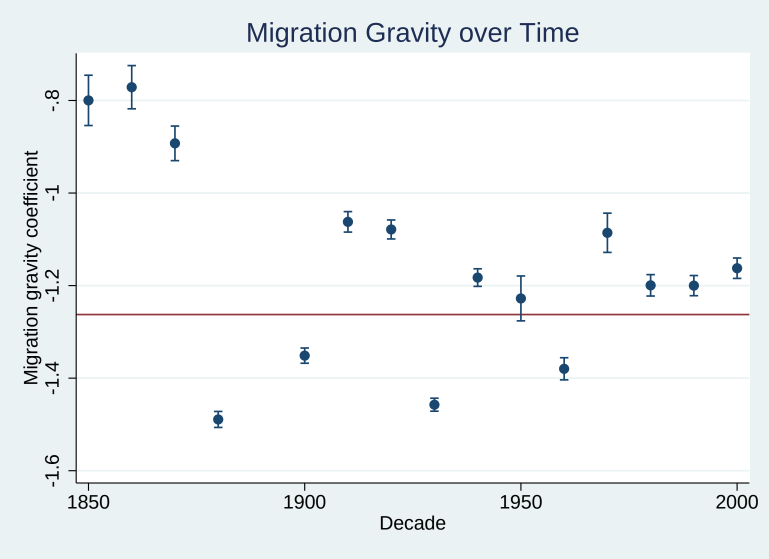
Estimation

- 6 key elasticities: $(\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta)$
- Estimation procedure with advantages important here:
  - Only need a little bilateral trade and migration data.
  - Simultaneously estimates the 6 elasticities.
  - Implemented using two 2SLS regressions:
    - Residuals from the regressions are $\bar{A}_{it}$ and $\bar{u}_{it}$.
- Step #1: Recover trade and migration costs from gravity equations:

$$\ln X_{ijt} = (1-\sigma)\,\kappa_t \ln dist_{ij} + \gamma_{it} + \delta_{jt} + \varepsilon_{ijt}$$

$$\ln L_{ijt} = -\theta\lambda_t \ln dist_{ij} + \rho_{it} + \pi_{jt} + \nu_{ijt}$$

Estimation Step #1: Recovering trade and migration costs



Estimation Step #2: Model inversion

- Define $T_{ij} \equiv \hat{\tau}_{ij}^{1-\sigma}$, $M_{ij} \equiv \hat{\mu}_{ij}^{-\theta}$, $p_{it} \equiv \frac{w_{it}}{A_{it}}$, and $Y_{it} \equiv w_{it} L_{it}$.
- Re-write equilibrium conditions as follows:
  - Goods market clearing:

$$p_{it}^{\sigma-1} = \sum_j T_{ij} \left(\frac{Y_{jt}}{Y_{it}} \right) P_{jt}^{\sigma-1}$$

$$P_{it}^{\sigma-1} = \sum_j T_{ji} \left(p_{jt}^{\sigma-1} \right)^{-1}$$

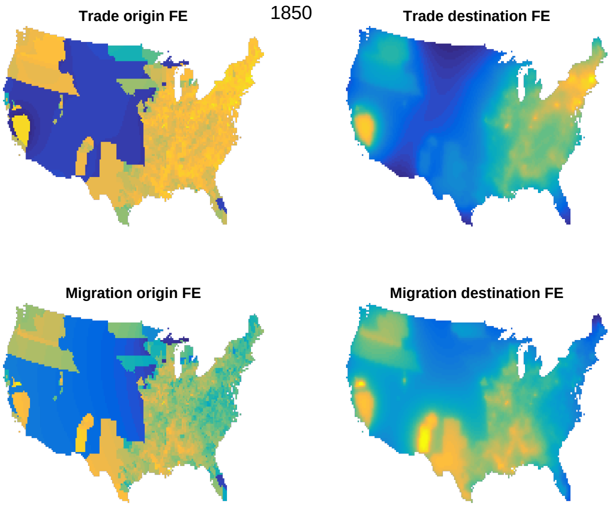
  - Labor market clearing:

$$\left(W_{it}^{\theta} \right)^{-1} = \sum_j M_{ji} \frac{L_{jt-1}}{L_{it}} \left(\Pi_{jt}^{\theta} \right)^{-1}$$

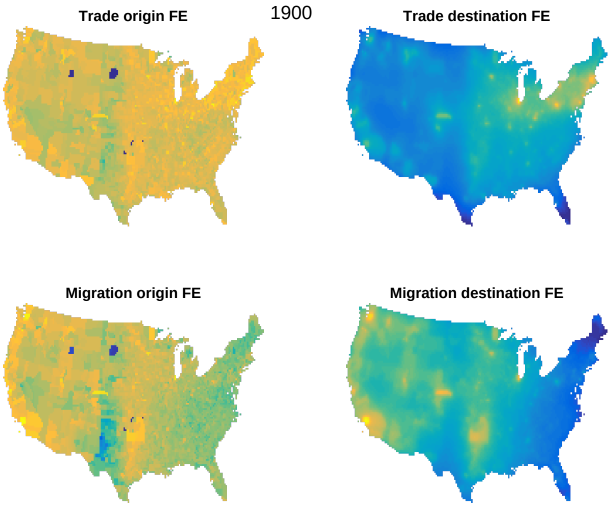
$$\Pi_{it}^{\theta} = \sum_i M_{ij} W_{jt}^{\theta}$$

- **Proposition 2**: *Given observed* $\{Y_{it}, L_{it}, L_{it-1}\}$, *there exists unique (to-scale)* $\{p_{it}^{\sigma-1}, P_{it}^{\sigma-1}, W_{it}^{\theta}, \Pi_{it}^{\theta}\}$.

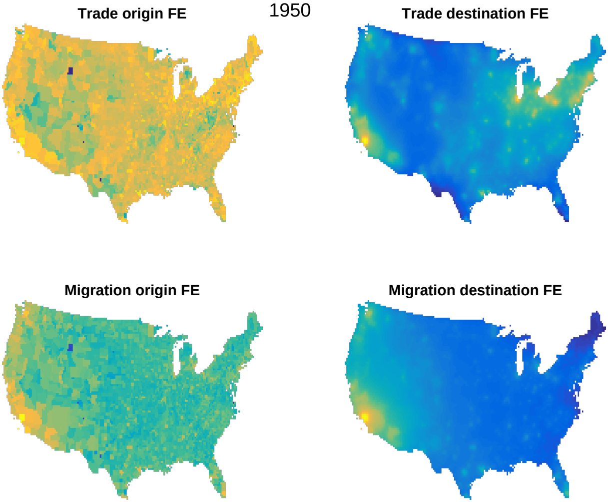
Model-inverted $\left\{p_{it}^{\sigma-1}, P_{it}^{\sigma-1}, W_{it}^{\theta}, \Pi_{it}^{\theta}\right\}$ values



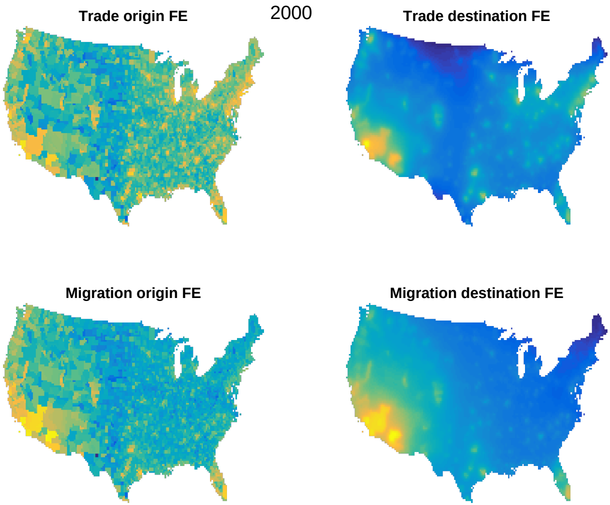
Model-inverted $\left\{p_{it}^{\sigma-1}, P_{it}^{\sigma-1}, W_{it}^{\theta}, \Pi_{it}^{\theta}\right\}$ values



Model-inverted $\left\{p_{it}^{\sigma-1}, P_{it}^{\sigma-1}, W_{it}^{\theta}, \Pi_{it}^{\theta}\right\}$ values



Model-inverted $\left\{p_{it}^{\sigma-1}, P_{it}^{\sigma-1}, W_{it}^{\theta}, \Pi_{it}^{\theta}\right\}$ values



Estimation Step #3: IV to recover model elasticities

- Step #3: Regress inverted $p_{it}^{\sigma-1}$ and W_{it}^{θ} on L_{it}, L_{it-1}, and w_{it}:

$$\ln\left(p_{it}^{\sigma-1}\right) = (\sigma-1)\ln w_{it} + \alpha_1(1-\sigma)\ln L_{it} + \alpha_2(1-\sigma)\ln L_{it-1} + (1-\sigma)\ln \bar{A}_{it}$$

$$\ln\left(W_{it}^{\theta}\right) = \theta\ln w_{it} + \left(\frac{1}{1-\sigma}\right)\ln\left(P_{it}^{1-\sigma}\right) + \beta_1\theta\ln L_{it} + \beta_2\theta\ln L_{it-1} + \theta\ln\bar{u}_{it}$$

- Issue: Residuals $\ln\bar{A}_{it}$ and $\ln\bar{u}_{it}$ correlated with endogenous outcomes.
  - Need instruments...

- For now: use IVs from equilibrium $\{\ln w_{it}, \ln L_{it}, \ln L_{it-1}\}$ in simulated model with:
  - Assume $\beta_1 = -0.3, \alpha_1 = \alpha_2 = \beta_2 = 0.1, \sigma = 9, \theta = 8,$
  - $\overline{A}_{it}$ and $\overline{u}_{it}$ proxied by geographic variables
  - Start model at $\{L_{i0}\}$ equal to observed 1800 population shares

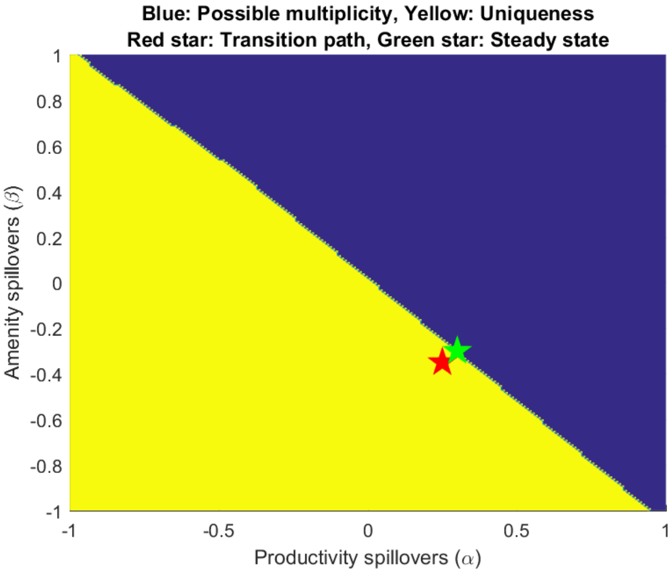
Table: Estimating elasticities and spillovers

	First stage			Second stage	
	(1)	(2)	(3)	(4)	(5)
	Wage	Pop.	Trade dest. FE	Trade orig. FE	Migr. dest. FE
Model log wage	0.612***				
	(0.110)				
Predicted log wage				-12.676***	11.736***
				(1.913)	(1.621)
Model log population		0.315***			
		(0.024)			
Predicted log population				3.034***	-4.000***
				(0.512)	(0.515)
Predicted log population 50 years ago				0.351***	-0.045*
				(0.028)	(0.025)
Model log price index			-4.141***		
			(0.052)		
Predicted log trade destination FE					0.240
					(0.187)
Elasticity of substitution σ				13.676***	49.821
				(1.913)	(36.513)
Migration elasticity θ					11.736***
					(1.621)
Contemporaneous productivity spillover α_1				0.239***	
				(0.010)	
Lagged productivity spillover α_2				0.028***	
				(0.003)	
Contemporaneous amenity spillover β_1					-0.341***
					(0.018)
Lagged amenity spillover β_2					-0.004*
					(0.002)
1800 Population	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes
Box-year FE	Yes	Yes	Yes	Yes	Yes
F-statistic	165.7	384.5	1298.0	299.0	255.2
R-squared	0.504	0.523	0.846	0.432	0.628
Observations	44408	44408	44408	44408	44408

Table: Estimating elasticities and spillovers over time

	Trade origin FE				Migration destination FE			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1850	1900	1950	2000	1850	1900	1950	2000
Elasticity of substitution σ	13.975***	14.421***	13.495***	16.216***	4.877***	4.247***	7.834**	4.505***
	(2.051)	(2.079)	(2.122)	(2.138)	(0.807)	(0.925)	(3.288)	(1.607)
Migration elasticity θ					8.449***	6.756***	5.582***	4.504***
					(1.678)	(1.699)	(1.733)	(1.744)
Contemporaneous productivity spillover α_1	0.144***	0.186***	0.266***	0.268***				
	(0.021)	(0.016)	(0.013)	(0.010)				
Lagged productivity spillover α_2	0.033***	0.032***	0.010	-0.010				
	(0.004)	(0.009)	(0.010)	(0.008)				
Contemporaneous amenity spillover β_1					-0.763***	-0.797***	-0.670***	-0.610***
					(0.104)	(0.142)	(0.136)	(0.148)
Lagged amenity spillover β_2					0.013**	0.048***	-0.071**	-0.232**
					(0.005)	(0.018)	(0.029)	(0.095)
1800 Population	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Box-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	196.8	196.8	196.8	196.8	179.0	179.0	179.0	179.0
Observations	0.441	0.441	0.441	0.441	0.638	0.638	0.638	0.638
N	44408	44408	44408	44408	44408	44408	44408	44408

Parameter Estimates



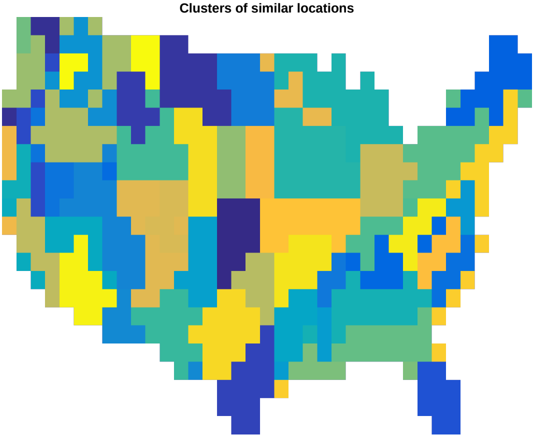
Quantifying Path Dependence

- Big picture question: How does history matter?

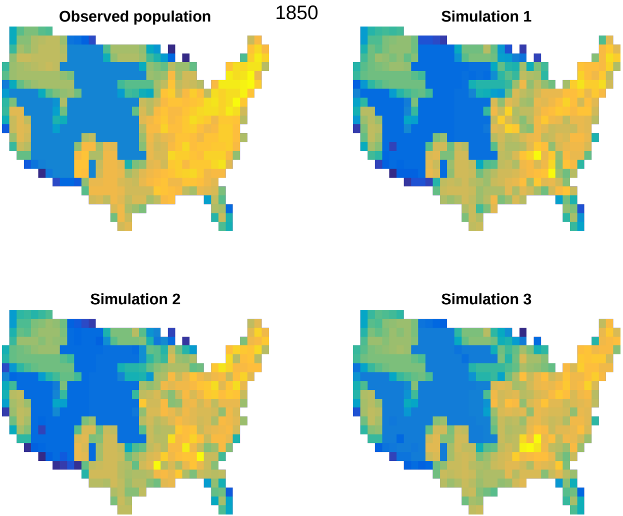
- AD (2018) simulate two types of counterfactual histories and compare those simulations with the factual history:
  1. Alternative starting points: random $\{L_{i0}\}$ for year 1800
  2. Alternative shocks along path: random $\overline{A}_{it}$ and $\overline{u}_{it}$ for years 1900 and 1950 (but hold 1850 and 2000 to factual values)

- In both cases, consider two questions:
  1. What happens to the *distribution* of economic activity?
  2. What happens to aggregate welfare (in eventual steady-state)?

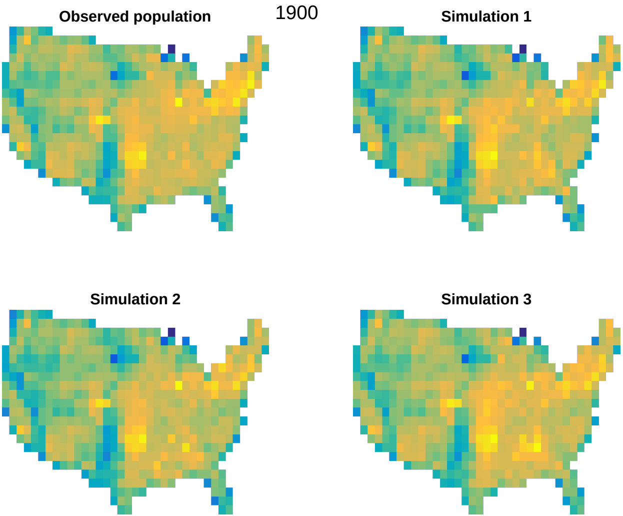
Simulations always redraw (without replacement) within geographic clusters



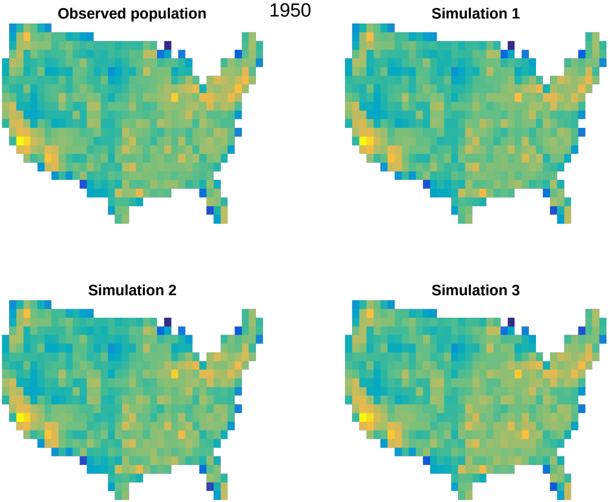
Shocking $\{L_{i0}\}$: examples from 3 simulations



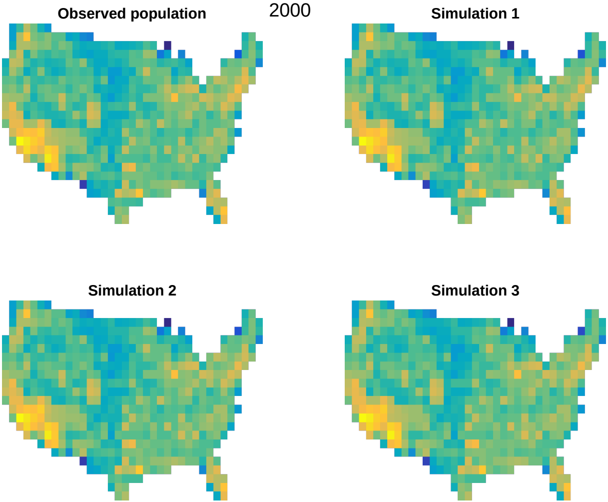
Shocking $\{L_{i0}\}$: examples from 3 simulations



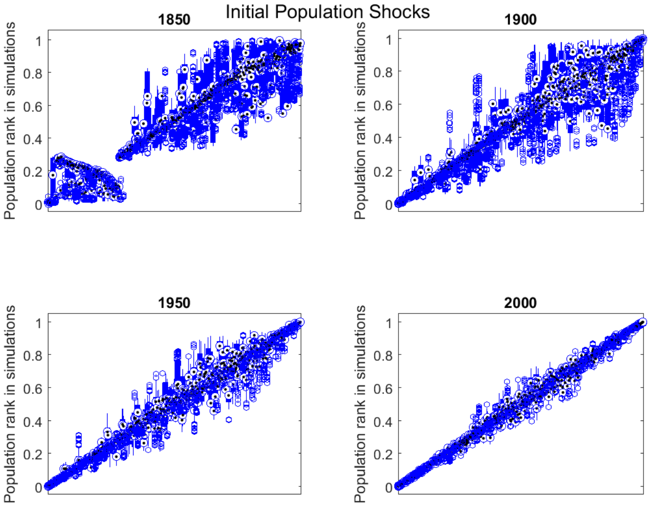
Shocking $\{L_{i0}\}$: examples from 3 simulations



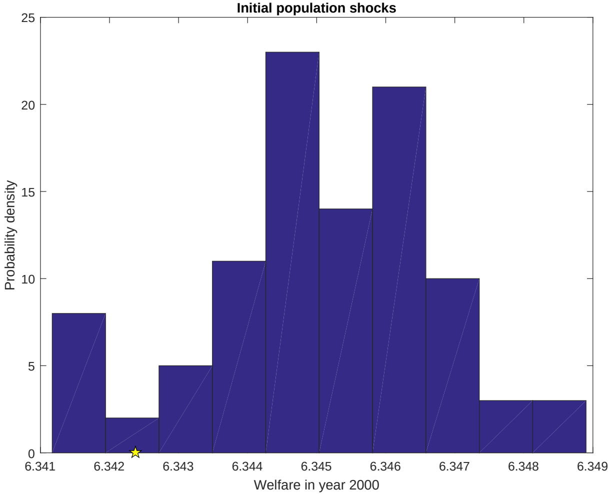
Shocking $\{L_{i0}\}$: examples from 3 simulations



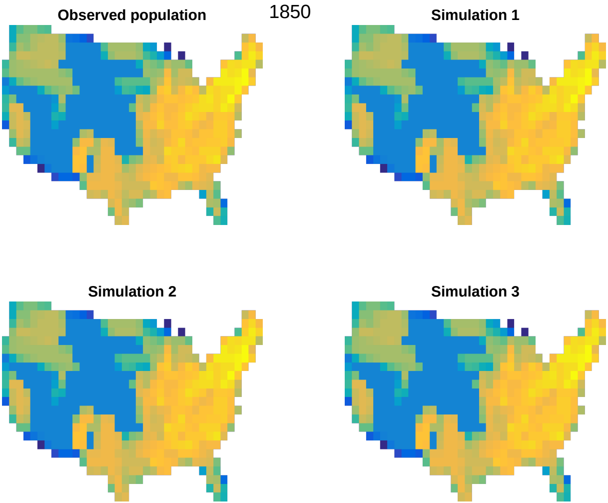
Shocking $\{L_{i0}\}$: effect on location of L_{it} across 200 simulations



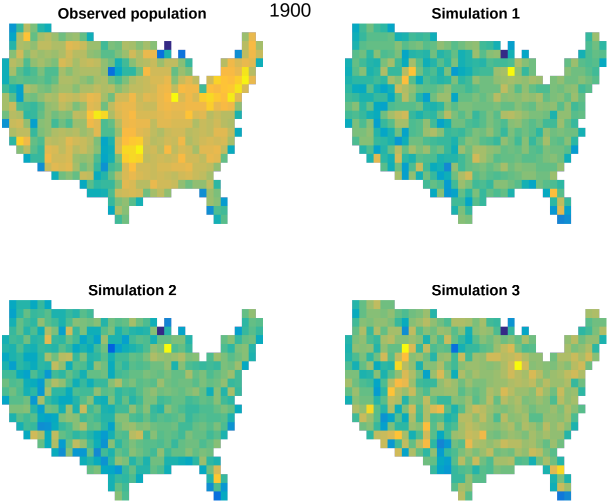
Shocking $\{L_{i0}\}$: effect on year 2000 welfare across 200 simulations



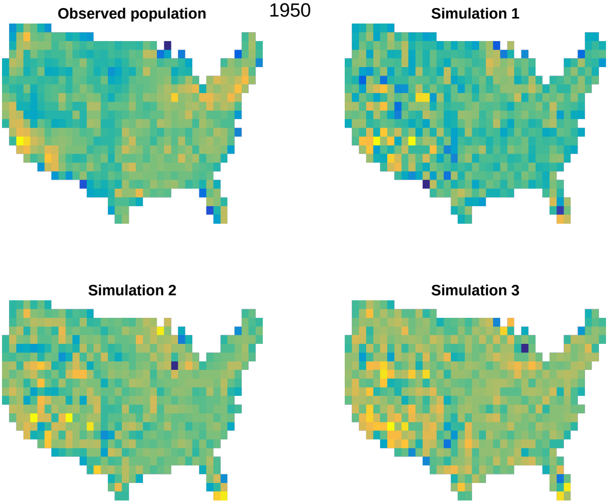
Shocking $\overline{A}_{it}$ in 1900 and 1950: examples from 3 simulations



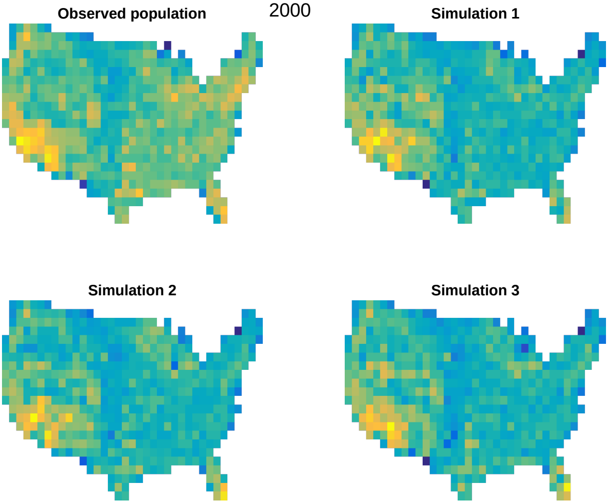
Shocking $\overline{A}_{it}$ in 1900 and 1950: examples from 3 simulations



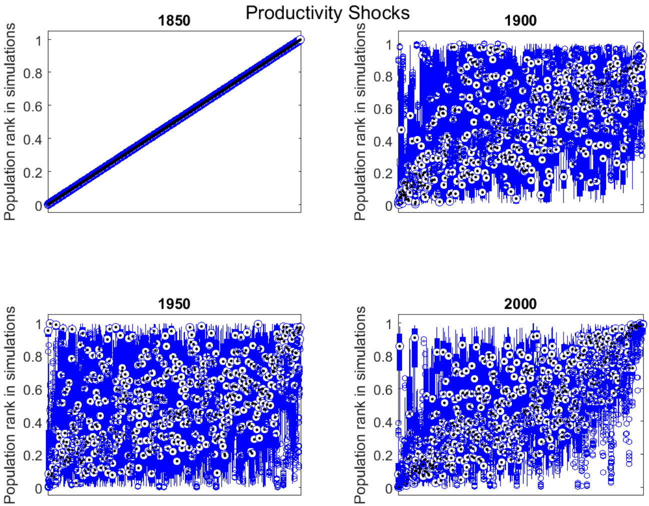
Shocking $\overline{A}_{it}$ in 1900 and 1950: examples from 3 simulations



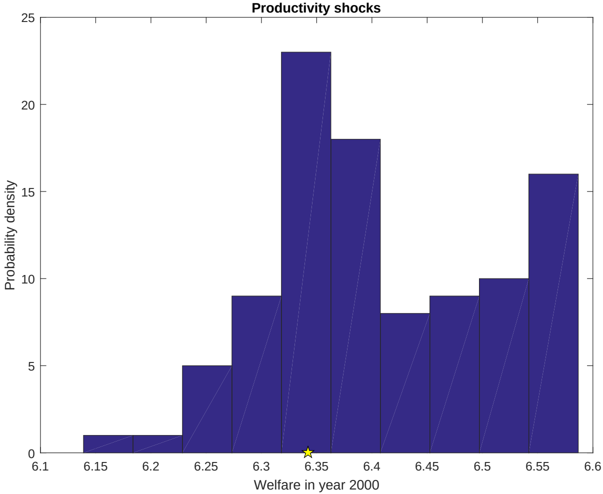
Shocking $\overline{A}_{it}$ in 1900 and 1950: examples from 3 simulations



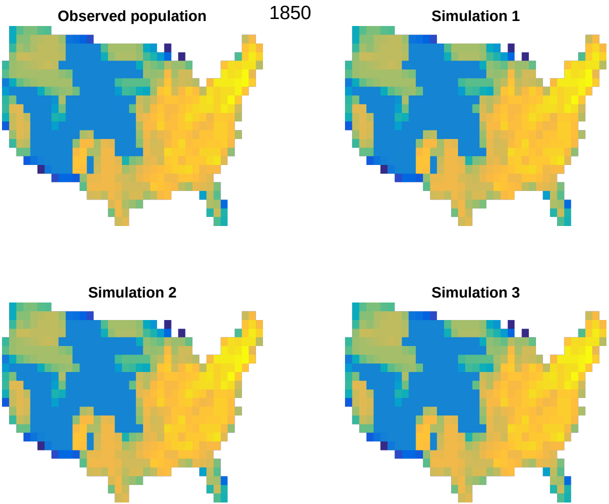
Shocking $\overline{A}_{it}$ in 1900 and 1950: effect on location of L_{it} across 200 simulations



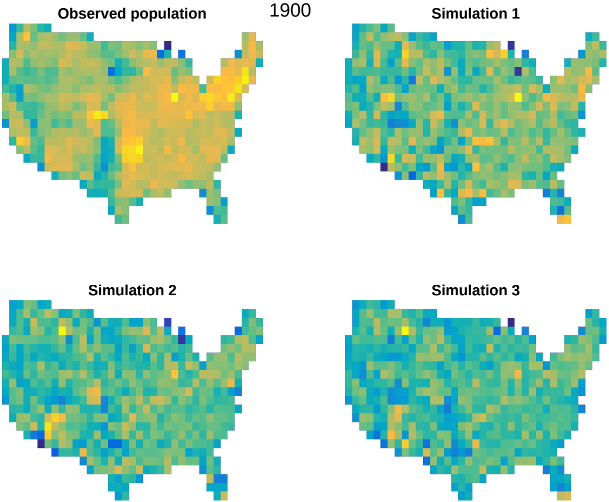
Shocking $\overline{A}_{it}$ in 1900 and 1950: effect on year 2000 welfare across 200 simulations



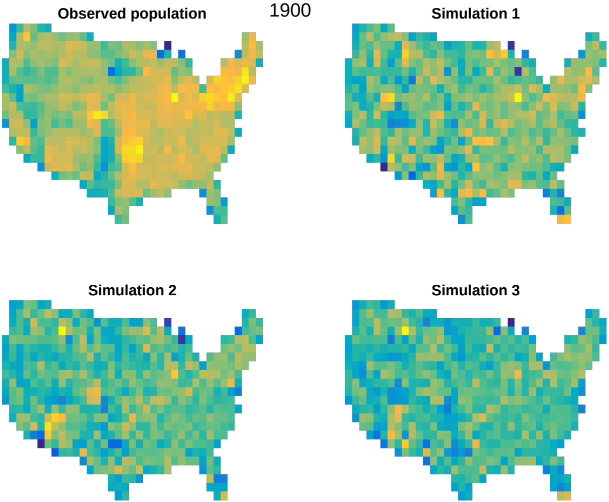
Shocking $\overline{u}_{it}$ in 1900 and 1950: examples from 3 simulations



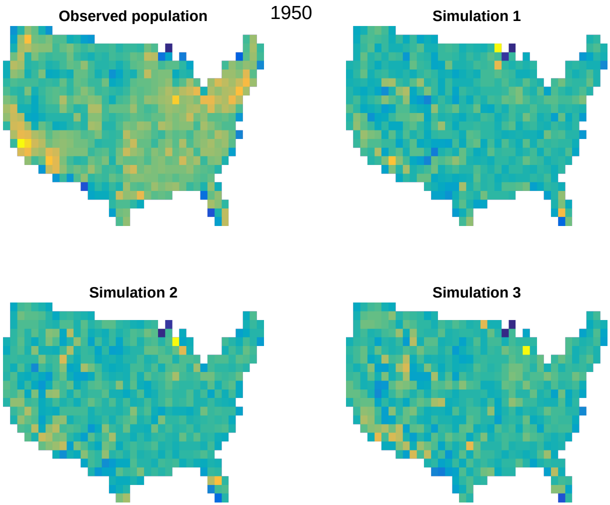
Shocking $\overline{u}_{it}$ in 1900 and 1950: examples from 3 simulations



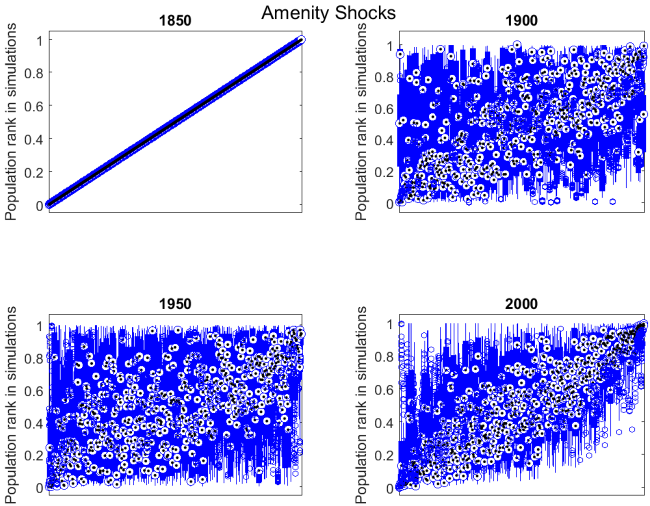
Shocking $\overline{u}_{it}$ in 1900 and 1950: examples from 3 simulations



Shocking $\overline{u}_{it}$ in 1900 and 1950: examples from 3 simulations



Shocking $\overline{u}_{it}$ in 1900 and 1950: effect on location of L_{it} across 200 simulations



Shocking $\overline{u}_{it}$ in 1900 and 1950: effect on year 2000 welfare across 200 simulations