2020 Lectures on Urban Economics

Lecture 6: Economic Geography and Path Dependence

Dave Donaldson (MIT)

16 July 2020
UEA Lectures 2020:
Economic Geography and Path Dependence
Dave Donaldson (MIT)
Economic activity is astoundingly concentrated over space—e.g. for the case of the U.S., 20% of GDP is produced in three cities (using only 1.5% of land area)

Why? (And, anyway, who cares?)

This lecture will discuss some answers.

Warning: my goal here is more to provide a “teaser” (examples—many of them narcissistic!—of the kinds of things that research has looked at, with different styles and vintages) than to survey the literature comprehensively. Some further reading mentioned at the end.
The US at Night
What about concentration of individual industries?

- No shortage of examples/anecdotes...
  - Marshall’s original examples (cutlery in Sheffield; jewelry in Birmingham)
  - Computers in Silicon Valley
  - Cars in Detroit
  - Entertainment in Las Vegas
  - Carpets in Dalton, Georgia
Holmes and Stevens (2004)

Figure 2:
Location of Large Manufacturing Plants (1947)

1 dot = plant 250+ employees

Figure 3:
Location of Large Manufacturing Plants (1999)

1 dot = plant 250+ employees
Figure 4: Location of Durum Wheat, Rice, Flue Tobacco, and Burley Tobacco

1 dot = 50,000 tons
Figure 5:
Location of Sugar Beet Plants and Sugar Beet Crops

- Red dot = 1,000 tons 1999 production
- Black dot = processing plant
Figure 6: Location of Anheuser-Busch Breweries and Population (2000)

- 1 dot = 10,000 people
- 1 dot = processing plant
Ellison and Glaeser (1997)

- EG (1997) aims to go beyond the anecdotes and ask just how concentrated is economic activity within any given industry in the US?

- Key point: What is the right null hypothesis?
  - If output, within an industry, is highly concentrated in a small number of plants, then that industry will look very concentrated spatially, simply by nature of the small number of plants. (Consider extreme case of one plant.)

- EG develop an index of localization that considers as its null hypothesis the random location of plants within an industry. They call this a “dartboard approach”.

- See also Duranton and Overman (2005) for statistical inference and corrections for the lumpiness of discrete spatial units.
EG (1997): Results ($\gamma = 0$ corresponds to $H_0$ of no excess agglomeration... “Slight concentration is remarkably widespread”)

**Fig. 1.**—Histogram of $\gamma$ (four-digit industries)

B. How Concentrated Are They?
In this subsection, we try to use our models to get a feel for how much concentration there is. We begin by imposing no structure across industries and simply computing the index $\gamma$ defined by (5) for each of the 459 four-digit industries in our sample. A complete list of the $\gamma$’s we find can be found in appendix C of Ellison and Glaeser (1994) and is also available from the authors on request.

A histogram illustrating the frequency distribution of these $\gamma$’s is presented in figure 1. In the figure, each bar represents the number of industries for which $\gamma$ lies in an interval of width 0.01. The distribution in the figure appears to be quite skewed, with the mean being 0.051 and the median being 0.026. The most striking feature of the figure is the large number of industries falling into the range we described as not very concentrated ($\gamma$, 0.02). The tallest bar is the one corresponding to values of $\gamma$ between zero and 0.01, and 43 percent of the industries have $\gamma$, 0.02. On the other side, the figure displays a thick right tail, with slightly more than a quarter of the

15 If one interprets $\gamma$’s as estimates of $\gamma_{na}$ $\gamma_{sa}$ $\gamma_{ns}$ $\gamma_{sa}$ (as opposed to estimates of the realized sum of squared differences between the $p$’s and the $x$’s), these $\gamma$’s are measured with substantial errors. To get a feel for the magnitudes, we computed standard errors by simulating a special case of our natural advantage model: that of Dirichlet-distributed state sizes. Among industries with $H$, 0.02, the mean of the estimated standard errors is 0.02. The means for industries with $H$ in the ranges 0.02±0.05, 0.05±0.10, and 0.10±1.0 are 0.024, 0.041, and 0.072, respectively.
EG (1997): Results

For industries that we might expect to be highly concentrated:
- Autos: $\gamma = 0.127$
- Auto parts: $\gamma = 0.089$
- Carpets (i.e. Dalton, GA): $\gamma = 0.378$
- Electronics (i.e. Silicon Valley): $\gamma = 0.059 - 0.142$

For industries that we might expect to not be highly concentrated:
- Bottled/canned soft drinks: $\gamma = 0.005$
- Newspaper: $\gamma = 0.002$
- Concrete: $\gamma = 0.012$
- Ice: $\gamma = 0.012$
### TABLE 4
**Most and Least Localized Industries**

<table>
<thead>
<tr>
<th>Four-Digit Industry</th>
<th>$H$</th>
<th>$G$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>15 Most Localized Industries</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2371 Fur goods</td>
<td>.007</td>
<td>.60</td>
<td>.63</td>
</tr>
<tr>
<td>2084 Wines, brandy, brandy spirits</td>
<td>.041</td>
<td>.48</td>
<td>.48</td>
</tr>
<tr>
<td>2252 Hosiery not elsewhere classified</td>
<td>.008</td>
<td>.42</td>
<td>.44</td>
</tr>
<tr>
<td>3533 Oil and gas field machinery</td>
<td>.015</td>
<td>.42</td>
<td>.43</td>
</tr>
<tr>
<td>2251 Women’s hosiery</td>
<td>.028</td>
<td>.40</td>
<td>.40</td>
</tr>
<tr>
<td>2273 Carpets and rugs</td>
<td>.013</td>
<td>.37</td>
<td>.38</td>
</tr>
<tr>
<td>2429 Special product sawmills not elsewhere classified</td>
<td>.009</td>
<td>.36</td>
<td>.37</td>
</tr>
<tr>
<td>3961 Costume jewelry</td>
<td>.017</td>
<td>.32</td>
<td>.32</td>
</tr>
<tr>
<td>2895 Carbon black</td>
<td>.054</td>
<td>.32</td>
<td>.30</td>
</tr>
<tr>
<td>3915 Jewelers’ materials, lapidary</td>
<td>.025</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>2874 Phosphatic fertilizers</td>
<td>.066</td>
<td>.32</td>
<td>.29</td>
</tr>
<tr>
<td>2061 Raw cane sugar</td>
<td>.038</td>
<td>.30</td>
<td>.29</td>
</tr>
<tr>
<td>2281 Yarn mills, except wool</td>
<td>.005</td>
<td>.27</td>
<td>.28</td>
</tr>
<tr>
<td>3534 Elevators and moving stairways</td>
<td>.030</td>
<td>.29</td>
<td>.28</td>
</tr>
<tr>
<td>3652 Prerecorded records and tapes</td>
<td>.070</td>
<td>.28</td>
<td>.25</td>
</tr>
<tr>
<td>2035 Pickles, sauces, salad dressings</td>
<td>.040</td>
<td>.27</td>
<td>.25</td>
</tr>
<tr>
<td>3482 Small-arms ammunition</td>
<td>.010</td>
<td>.25</td>
<td>.25</td>
</tr>
<tr>
<td>3524 Steel investment foundries</td>
<td>.040</td>
<td>.25</td>
<td>.25</td>
</tr>
<tr>
<td>3534 Elevators and moving stairways</td>
<td>.030</td>
<td>.29</td>
<td>.28</td>
</tr>
<tr>
<td>2052 Cookies and crackers</td>
<td>.030</td>
<td>.29</td>
<td>.28</td>
</tr>
<tr>
<td>2098 Macaroni and spaghetti</td>
<td>.030</td>
<td>.29</td>
<td>.28</td>
</tr>
<tr>
<td>3262 Vitreous china table, kitchenware</td>
<td>.13</td>
<td>.12</td>
<td>.003</td>
</tr>
<tr>
<td>2035 Pickles, sauces, salad dressings</td>
<td>.01</td>
<td>.01</td>
<td>.003</td>
</tr>
<tr>
<td>3821 Laboratory apparatus and furniture</td>
<td>.02</td>
<td>.02</td>
<td>.002</td>
</tr>
<tr>
<td>2062 Cane sugar refining</td>
<td>.11</td>
<td>.10</td>
<td>.002</td>
</tr>
<tr>
<td>3433 Heating equipment except electric</td>
<td>.01</td>
<td>.01</td>
<td>.002</td>
</tr>
<tr>
<td><strong>15 Least Localized Industries</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3021 Rubber and plastics footwear</td>
<td>.06</td>
<td>.05</td>
<td>-.013</td>
</tr>
<tr>
<td>2032 Canned specialties</td>
<td>.03</td>
<td>.02</td>
<td>-.012</td>
</tr>
<tr>
<td>2082 Malt beverages</td>
<td>.04</td>
<td>.03</td>
<td>-.010</td>
</tr>
<tr>
<td>3635 Household vacuum cleaners</td>
<td>.18</td>
<td>.17</td>
<td>-.009</td>
</tr>
<tr>
<td>3652 Prerecorded records and tapes</td>
<td>.04</td>
<td>.03</td>
<td>-.008</td>
</tr>
<tr>
<td>3482 Small-arms ammunition</td>
<td>.18</td>
<td>.17</td>
<td>-.004</td>
</tr>
<tr>
<td>3324 Steel investment foundries</td>
<td>.04</td>
<td>.04</td>
<td>-.003</td>
</tr>
<tr>
<td>3534 Elevators and moving stairways</td>
<td>.03</td>
<td>.03</td>
<td>-.001</td>
</tr>
<tr>
<td>2052 Cookies and crackers</td>
<td>.03</td>
<td>.03</td>
<td>-.009</td>
</tr>
<tr>
<td>2098 Macaroni and spaghetti</td>
<td>.03</td>
<td>.03</td>
<td>-.008</td>
</tr>
<tr>
<td>3262 Vitreous china table, kitchenware</td>
<td>.13</td>
<td>.12</td>
<td>-.006</td>
</tr>
<tr>
<td>2035 Pickles, sauces, salad dressings</td>
<td>.01</td>
<td>.01</td>
<td>-.003</td>
</tr>
<tr>
<td>3821 Laboratory apparatus and furniture</td>
<td>.02</td>
<td>.02</td>
<td>-.002</td>
</tr>
<tr>
<td>2062 Cane sugar refining</td>
<td>.11</td>
<td>.10</td>
<td>.002</td>
</tr>
<tr>
<td>3433 Heating equipment except electric</td>
<td>.01</td>
<td>.01</td>
<td>.002</td>
</tr>
</tbody>
</table>

If firms choose identical locations, with natural advantages being independent across geographic areas. If, on the other hand, the effect of spillovers (or the spatial correlation of natural advantage) is smoothly declining with distance, then those $\gamma$'s will reflect the excess probability with which pairs of firms tend to locate in the same county, state, and region, respectively. To investigate the geographic scope of spillovers, we estimated $\gamma$'s from our county/three-digit data set using counties, states, and the nine census regions as the units of observation.
EG (1999): Extend the EG (1997) method with empirical estimates of role played by natural advantage

Table 1—Effect of “Natural Advantages”
on State-Industry Employment

<table>
<thead>
<tr>
<th>A. State variable × industry variable</th>
<th>Coefficient (t statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Electricity price × electricity use</td>
<td>0.170 (17.62)</td>
</tr>
<tr>
<td>(b) Natural gas price × natural gas use</td>
<td>0.117 (6.91)</td>
</tr>
<tr>
<td>(c) Coal price × coal use</td>
<td>0.119 (4.55)</td>
</tr>
<tr>
<td>(d) Percentage farmland × agricultural inputs</td>
<td>0.026 (2.58)</td>
</tr>
<tr>
<td>(e) Per capita cattle × livestock inputs</td>
<td>0.053 (5.08)</td>
</tr>
<tr>
<td>(f) Percentage timberland × lumber inputs</td>
<td>0.152 (11.98)</td>
</tr>
<tr>
<td>(g) Average mfg wage × wages/value added</td>
<td>0.059 (4.11)</td>
</tr>
<tr>
<td>(h) Average mfg wage × exports/output</td>
<td>−0.014 (−1.28)</td>
</tr>
<tr>
<td>(i) Average mfg wage × import competition</td>
<td>0.036 (3.10)</td>
</tr>
<tr>
<td>(j) Percentage without HS degree × percentage unskilled</td>
<td>0.157 (7.38)</td>
</tr>
<tr>
<td>(k) Unionization percentage × percentage precision products</td>
<td>0.100 (12.17)</td>
</tr>
<tr>
<td>(l) Percentage with B.A. or more × percentage executive/professional</td>
<td>0.170 (12.70)</td>
</tr>
<tr>
<td>(m) Coast dummy × heavy exports</td>
<td>−0.031 (−2.20)</td>
</tr>
<tr>
<td>(n) Coast dummy × heavy imports</td>
<td>0.017 (0.92)</td>
</tr>
<tr>
<td>(o) Population density × percentage to consumers</td>
<td>0.043 (3.68)</td>
</tr>
<tr>
<td>(p) (Income share − mfg share) × percentage to consumers</td>
<td>0.025 (4.49)</td>
</tr>
</tbody>
</table>

Note: Letters in this column refer to state and industry variables in part A of the table.
EG (1999): Extend the EG (1997) method with empirical estimates of role played by natural advantage

**Table 2—Estimates of Residual Geographic Concentration after Accounting for Observed Natural Advantages**

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean $\gamma$</th>
<th>Percentage of industries with $\gamma$ in range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\gamma$</td>
<td>&lt;0.0</td>
</tr>
<tr>
<td>A</td>
<td>0.051</td>
<td>2.8</td>
</tr>
<tr>
<td>B</td>
<td>0.048</td>
<td>3.9</td>
</tr>
<tr>
<td>C</td>
<td>0.045</td>
<td>3.1</td>
</tr>
<tr>
<td>D</td>
<td>0.041</td>
<td>4.4</td>
</tr>
</tbody>
</table>

Notes: Models A–D are different models of natural advantage: (A) no cost variables; (B) cost interactions introduced; (C) cost interactions plus dummies for two-digit industries; (D) cost interactions plus dummies for three-digit industries.
Why is output so agglomerated?

Three broad explanations:

1. Some production input is exogenously agglomerated.
   - Natural resources and other immobile factors (such as, plausibly, for case of the wine industry in EG (1997))

2. Some consumption amenity is exogenously agglomerated
   - Nice places to live (for place-based amenities that are non-tradable)

3. Some production input agglomerates endogenously
   - Positive externality (i.e. spillover) in production that depends on proximity. Anecdotal/case-study accounts suggest that this very likely explains Silicon Valley, Detroit, Boston biotech, carpets in Dalton, etc.
   - Similarly, can have positive externality in amenities
Why is output so agglomerated?

- This points are often simplified into:

  1. “First-nature geographies”: exogenous locational characteristics that affect productivity and amenities
  2. “Second-nature geographies”: endogenous locational characteristics that affect productivity and amenities

- We will now dig into models that feature both of these sources
Consider first (though only very briefly) Paul Krugman’s famous “core-periphery” model of economic geography:

- 2 symmetric locations (so no first-nature geographies at all)
- 2 sectors:
  - A: CRTS, homogenous good, perfect competition
  - M: Dixit-Stiglitz monopolistic competition with CES ($\sigma$)—isomorphic to model with aggregate (within sector M) external economies of scale of elasticity $1/(\sigma - 1)$.

- “Iceberg” transport costs $\tau$ (and let $\varphi \equiv \tau^{1-\sigma}$)

- Workers (single factor of production):
  - Freely mobile across industries
  - Fixed fraction of workers are immobile across locations
  - Other workers are freely mobile across locations—seek location with highest real wage
Krugman (1991): Equilibria

Figure 11.1 The Tomahawk diagram

© Brakman, Garretsen, and van Marrewijk, 2008
Evidence for agglomeration externalities (source of second-nature geographies) seems strong:

- Case studies (e.g. Silicon Valley)
- Some recent direct estimates (e.g. Ahlfeldt et al (2015) on Berlin Wall; Greenstone, Hornbeck and Moretti (2010) on Million Dollar Plants; Kline and Moretti (2014) on Tennessee Valley Authority)

Long theoretical tradition highlights very important implications:

- Potential for multiple equilibria in static models
- In dynamic models, potential for:
  - Multiple stable steady-states
  - Hence potential for path dependence—i.e. initial conditions, or long-redundant shocks, still matter for outcomes today
  - Cheap policies to promote movement to better steady-state

But is there direct empirical evidence for such path dependence?
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:
  - ...

- 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.
wartime growth should asymptotically approach unity as the end period increases. In the last column of Table 3 we repeat the regression, only now extending the endpoint to 1965 instead of 1960. The estimated coefficient now reaches -1.027. That is, after controlling for prewar growth trends, by 1965 cities have entirely reversed the damage due to the war. Again, the impact of reconstruction subsidies also lessens as we move into the future. Together, these results suggest that the effect of the temporary shocks vanishes completely in less than 20 years.

One possible objection to our interpretation is that in most cases, the population changes responded much more to refugees than deaths. Of the 144 cities with positive casualties, the average number of deaths per capita was only 1 percent. Most of the population movement that we observe in our data is due to the fact that the vast destruction of buildings forced people to live elsewhere. However, forcing them to move out of their cities for a number of years may not have sufficed to overcome the social networks and other draws of their home cities. Hence it may seem uncertain whether they are moving back to take advantage of particular characteristics of these locations or simply moving back to the only real home they have known.

However, there are two cases in which this argument cannot be made: Hiroshima and Nagasaki. In those cities, the number of deaths was such that if these cities recovered their populations, it could not be because residents who temporarily moved out of the city returned in subsequent years. We have already noted that our data underestimates casualties in these cities. Even so, our data suggest that the nuclear bombs immediately killed 8.5 percent of Nagasaki's population and 20.8 percent of Hiroshima's population. Moreover, given that many Japanese were worried about radiation poisoning and actively discriminated against atomic bomb victims, it is unlikely that residents felt an unusually strong attachment to these cities or that other Japanese felt a strong desire to move there. Another reason why these cities are interesting to consider is that they were not particularly large or famous cities in Japan. Their 1940 populations made them the 8th and 12th largest cities in Japan. Both cities were close to other cities of comparable size so that it would have been relatively easy for other cities to absorb the populations of these devastated cities.

In Figure 2 we plot the population of these two cities. What is striking in the graph is that even in these two cities there is a clear indication that they returned to their prewar growth trends. This process seems to have taken a little longer in Hiroshima than in other cities, but this is not surprising given the level of destruction.
Sit+1 = Sit + 12it*I*

If \( p \in [0, 1) \), then city share is stationary and any shock will dissipate over time. In other words, these two hypotheses can be distinguished by identifying the parameter \( p \).

One approach to investigating the magnitude of \( p \) is to search for a unit root. It is well known that unit root tests usually have little power to separate \( p < 1 \) from \( p = 1 \). This is due to the fact that in traditional unit root tests the innovations are not observable and so identify \( p \) with very large standard errors. A major advantage of our data set is that we can easily identify the innovations due to bombing. In particular, since by hypothesis the innovation, \( v_{it} \), is uncorrelated with the error term (in square brackets), then if we can identify the innovation, we can obtain an unbiased estimate of \( p \).

An obvious method of looking at the innovation is to use the growth rate from 1940 to 1947. However, this measure of the innovation may contain not only information about the bombing but also past growth rates. This is a measurement error problem that could bias our estimates in either direction depending on \( p \). In order to solve this, we instrument the growth rate from 1940-1947 with buildings destroyed per capita and deaths per capita.  

We can obtain a feel for the data by considering the impact of bombing on city growth rates. As we argued earlier, if city growth rates follow a random walk, then all shocks to cities should be permanent. In this case, one should expect to see no relationship between historical shocks and future growth rates. Moreover, if one believes that there is positive serial correlation in the data, then one should expect to see a positive correlation between past and future growth rates. By contrast, if one believes that location-specific factors are crucial in understanding the distribution of population, then one should expect to see a negative relationship between a historical shock and the subsequent growth rate. In Figure 1 we present a plot of the growth rate between 1940 and 1947 or \( S_{1947} - S_{1940} = \gamma_{1947} + [P_{1947} - S_{1940}] \).

Our measure of the innovation is the growth rate between 1940 and 1947 or \( S_{1947} - S_{1940} = \gamma_{1947} + [P_{1947} - S_{1940}] \). This is clearly correlated with the error term in the estimating equation, hence we instrument.

Figure 1. Effects of Bombing on Cities with More than 30,000 Inhabitants

Note: The figure presents data for cities with positive casualty rates only.

In Table 2, we present a regression showing the power of our instruments. Deaths per capita and destruction per capita explain about 41 percent of the variance in population growth of cities between 1940 and 1947. Interestingly, although both have the expected signs, destruction seems to have had a more pronounced effect on the populations of cities. Presumably, this is because, with a few notable exceptions, the number of people killed was only a few percent of the city's population.

We now turn to test whether the temporary shocks give rise to permanent effects. In order to estimate equation (4), we regress the growth rate of cities between 1947 and 1960 on the growth rate between 1940 and 1947 using deaths and destruction per capita as instruments for the wartime growth rates. The coefficient on growth between 1940 and 1947 corresponds to \( (p - 1) \). In addition, we include government subsidies to cities to control for policies designed to rebuild cities.
Table 1
Evolution of Japanese manufacturing during World War II
(Quantum Indices from Japanese Economic Statistics)

<table>
<thead>
<tr>
<th>Industry</th>
<th>1941</th>
<th>1946</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>206.2</td>
<td>27.4</td>
<td>-87%</td>
</tr>
<tr>
<td>Machinery</td>
<td>639.2</td>
<td>38.0</td>
<td>-94%</td>
</tr>
<tr>
<td>Metals</td>
<td>270.2</td>
<td>20.5</td>
<td>-92%</td>
</tr>
<tr>
<td>Chemicals</td>
<td>252.9</td>
<td>36.9</td>
<td>-85%</td>
</tr>
<tr>
<td>Textiles and Apparel</td>
<td>79.4</td>
<td>13.5</td>
<td>-83%</td>
</tr>
<tr>
<td>Processed Food</td>
<td>89.9</td>
<td>54.2</td>
<td>-40%</td>
</tr>
<tr>
<td>Printing and Publishing</td>
<td>133.5</td>
<td>32.7</td>
<td>-76%</td>
</tr>
<tr>
<td>Lumber and Wood</td>
<td>187.0</td>
<td>91.6</td>
<td>-51%</td>
</tr>
<tr>
<td>Stone, Clay, Glass</td>
<td>124.6</td>
<td>29.4</td>
<td>-76%</td>
</tr>
</tbody>
</table>
Table 1

Evolution of Japanese manufacturing during World War II (Quantum Indices from Japanese Economic Statistics)

<table>
<thead>
<tr>
<th></th>
<th>1941</th>
<th>1946</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry</td>
<td>206.2</td>
<td>27.4</td>
<td>-87%</td>
</tr>
<tr>
<td>Machinery</td>
<td>639.2</td>
<td>38.0</td>
<td>-94%</td>
</tr>
<tr>
<td>Metals</td>
<td>270.2</td>
<td>20.5</td>
<td>-92%</td>
</tr>
<tr>
<td>Chemicals</td>
<td>252.9</td>
<td>36.9</td>
<td>-85%</td>
</tr>
<tr>
<td>Textiles and Apparel</td>
<td>79.4</td>
<td>13.5</td>
<td>-83%</td>
</tr>
<tr>
<td>Processed Food</td>
<td>89.9</td>
<td>54.2</td>
<td>-40%</td>
</tr>
<tr>
<td>Printing and Publishing</td>
<td>133.5</td>
<td>32.7</td>
<td>-76%</td>
</tr>
<tr>
<td>Lumber</td>
<td>187.0</td>
<td>91.6</td>
<td>-51%</td>
</tr>
<tr>
<td>Stone, Clay, Glass</td>
<td>124.6</td>
<td>29.4</td>
<td>-76%</td>
</tr>
</tbody>
</table>

Table 2

Correlation of Growth Rates of Industries Within Cities 1938 to 1948

<table>
<thead>
<tr>
<th></th>
<th>Machinery</th>
<th>Metals</th>
<th>Chemicals</th>
<th>Textiles</th>
<th>Food</th>
<th>Printing</th>
<th>Lumber</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metals</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemicals</td>
<td>0.30</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Textiles</td>
<td>0.12</td>
<td>0.35</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food</td>
<td>0.32</td>
<td>0.65</td>
<td>0.31</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Printing</td>
<td>0.11</td>
<td>0.30</td>
<td>0.04</td>
<td>0.29</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lumber</td>
<td>0.23</td>
<td>0.35</td>
<td>0.21</td>
<td>0.25</td>
<td>0.25</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>Ceramics</td>
<td>0.13</td>
<td>0.53</td>
<td>0.36</td>
<td>0.38</td>
<td>0.50</td>
<td>0.41</td>
<td>0.23</td>
</tr>
<tr>
<td>Asset Category</td>
<td>Decline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>25.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buildings</td>
<td>24.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harbors and canals</td>
<td>7.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bridges</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrial machinery and equipment</td>
<td>34.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Railroads and tramways</td>
<td>7.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cars</td>
<td>21.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ships</td>
<td>80.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electric power generation facilities</td>
<td>10.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telecommunication facilities</td>
<td>14.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water and sewerage works</td>
<td>16.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observations we find the same kind of mean reversion as Davis and Weinstein (2002, Figure 1) found in city population data.

Regression Results

In this section, we present our threshold regression results. Because it is possible that multiple equilibria arise at one level of aggregation even if not at another, we consider this at various levels of aggregation. We consider it first using the city population data considered in Davis and Weinstein (2002). The analysis of that data is augmented here by our new approach which sharpens the contrast between the theory of unique versus multiple equilibria and which also places the theories on a more even footing in our estimation approach. Thereafter, we consider the same questions using data on city aggregate manufacturing and city-industry observations for eight manufacturing industries. Since manufacturing is less than half of all economic activity within a typical city, it should be clear that even if population in a city were to recover from the shocks, this need not be true of aggregate city-manufacturing. The same point holds a fortiori for particular industries within manufacturing, which we also examine.

We begin by considering city population data. Column 1 of Table 4 replicates the Davis and Weinstein (2002) results using population data. The IV estimate in column 1 tests a null of a unique stable equilibrium by asking if we can reject $C^\circledast$ Blackwell Publishing, Inc. 2008.
that the coefficient on the wartime (1940–1947) growth rate is minus unity. We cannot reject a coefficient of minus unity, hence cannot reject a null that there is a unique stable equilibrium. We also find that regionally-directed government reconstruction expenses following the war had no significant impact on city sizes 20 years after the war.

We next apply our threshold regression approach described above to testing for multiple equilibria. This places unique and multiple equilibria on an even footing. The results are reported in the remaining columns of Table 4. In column 2 of Table 4, we present the results for the estimation of equation (11) in the case in which there is a unique equilibrium. Given how close our previous estimate of \( \theta \) was to 0 (minus unity on wartime growth), it is not surprising that the estimates of the other parameters do not change much when we constrain \( \theta \) to take on this value.

Columns 3–5 present the results for threshold regressions premised on various numbers of equilibria. In principle, we could have considered the possibility of more than four equilibria. However, neither the data plots nor any of the regression results suggested that raising the number of potential equilibria was likely to improve the results.

FIGURE 8: Prewar vs Postwar Growth Rate.
Bleakley and Lin (QJE, 2012)

- BL (2012) look for an event that diminished a location’s natural (i.e. exogenous) productivity or amenity advantage over other locations. [For other examples of this type of reasoning, see: Redding, Sturm and Wolf (2011); Henderson, Storeyard and Weil (2018); and Michaels and Rauch (2018)]

- 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’, a 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)
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.
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.
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.
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).
TABLE II

UPSTREAM WATERSHED AND CONTEMPORARY POPULATION DENSITY

<table>
<thead>
<tr>
<th>Specifications:</th>
<th>Basic</th>
<th>Other spatial controls</th>
<th>Water power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance from various features</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State fixed effects</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Explanatory variables:

Panel A: Census Tracts, 2000, N = 21452
Portage site times 0.467 0.467 0.500 0.496 0.452
upstream watershed (0.175)** (0.164)** (0.114)** (0.173)** (0.177)**
Binary indicator 1.096 1.000 1.111 1.099 1.056
for portage site (0.348)** (0.326)** (0.219)** (0.350)** (0.364)**
Portage site times
horsepower/100k
I(horsepower > 2000)

Panel B: Nighttime Lights, 1996-97, N = 65000
Portage site times 0.418 0.352 0.456 0.415 0.393
upstream watershed (0.115)** (0.102)** (0.113)** (0.116)** (0.111)**
Binary indicator 0.463 0.424 0.421 0.462 0.368
for portage site (0.116)** (0.111)** (0.121)** (0.116)** (0.132)**
Portage site times
horsepower/100k
I(horsepower > 2000)

Panel C: Counties, 2000, N = 3480
Portage site times 0.443 0.372 0.423 0.462 0.328
upstream watershed (0.209)** (0.185)** (0.207)** (0.215)** (0.154)**
Binary indicator for portage site 0.890 0.834 0.742 0.889 0.587
Portage site times
horsepower/100k
I(horsepower > 2000)
### TABLE III

**Proximity to Historical Portage Site and Historical Factors**

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

**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>
<thead>
<tr>
<th></th>
<th>Panel A. Portage and contemporary factors</th>
<th>Panel B. Portage and contemporary factors, conditioned on contemporary density</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(1)</td>
</tr>
<tr>
<td>Housing units, 1990</td>
<td>0.910**</td>
<td>0.005**</td>
</tr>
<tr>
<td>Median rents, 1990</td>
<td>0.110**</td>
<td>0.014**</td>
</tr>
<tr>
<td>Median values, 1990</td>
<td>0.108**</td>
<td>–0.001</td>
</tr>
<tr>
<td>Interstates, 2000</td>
<td>0.602**</td>
<td>0.159</td>
</tr>
<tr>
<td>Major roads, 2000</td>
<td>0.187**</td>
<td>–0.064</td>
</tr>
<tr>
<td>Rail, 2000</td>
<td>0.858**</td>
<td>0.182</td>
</tr>
<tr>
<td>Travel time to work, 1990</td>
<td>–0.554**</td>
<td>–0.447</td>
</tr>
<tr>
<td>Crime, 1995</td>
<td>1.224***</td>
<td>–0.007</td>
</tr>
<tr>
<td>Born in state, 1990</td>
<td>0.832***</td>
<td>–0.025</td>
</tr>
<tr>
<td>Water use, 1995</td>
<td>0.549***</td>
<td>–0.153</td>
</tr>
<tr>
<td>Federal expend., 1997</td>
<td>1.063***</td>
<td>0.032</td>
</tr>
<tr>
<td>Gov't. empl., 1997</td>
<td>1.001***</td>
<td>0.114</td>
</tr>
</tbody>
</table>

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.

Table IV

<table>
<thead>
<tr>
<th>Explanatory variables:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy for proximity to portage site</td>
<td>0.910***</td>
<td>0.110***</td>
<td>0.108**</td>
<td>0.602**</td>
<td>0.187**</td>
<td>0.858**</td>
<td>–0.554***</td>
<td>1.224***</td>
<td>0.832***</td>
<td>0.549***</td>
<td>1.063***</td>
<td>1.001***</td>
</tr>
<tr>
<td></td>
<td>(0.243)**</td>
<td>(0.040)**</td>
<td>(0.053)**</td>
<td>(0.228)**</td>
<td>(0.071)**</td>
<td>(0.177)**</td>
<td>(0.492)**</td>
<td>(0.318)**</td>
<td>(0.186)**</td>
<td>(0.197)**</td>
<td>(0.343)**</td>
<td>(0.283)**</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log hourly wage</td>
<td>OLS</td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
</tr>
<tr>
<td>Log population density</td>
<td>0.049</td>
<td>0.085</td>
<td>0.089</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>(0.003)**</td>
<td>(0.032)**</td>
<td>(0.030)**</td>
<td>(0.028)**</td>
</tr>
</tbody>
</table>

**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 NR² 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>
<thead>
<tr>
<th>TABLE VI</th>
<th>INTERACTION OF HISTORICAL FACTORS WITH GROWTH AT PORTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Baseline estimate</td>
<td>(2) Warm climate</td>
</tr>
<tr>
<td>Dummy for proximity to portage site × 20th century</td>
<td>0.456</td>
</tr>
<tr>
<td>Additional factor (column heading) × 20th century</td>
<td>0.124</td>
</tr>
<tr>
<td>Dummy for portage × add'l factor × 20th century</td>
<td>−0.402</td>
</tr>
</tbody>
</table>

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.
2020 Lectures on Urban Economics

Short Break – We are back in a few minutes
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 setup: Geography

- $N$ locations. Each location $i \in \{1, \ldots, N\}$ in each time period $t \in \{1, \ldots\}$ 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 $\mu_{ijt} \geq 1$. 

Donaldson (MIT) 
UEA Lectures 2020
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 $t$.

- All initial populations $\{L_{i0}\}$ given exogenously.
Model setup: Production and Consumption

- **Production**
  - Perfect competition, (adult) labor only factor of production. Quantity produced:
    \[
    Q_{it} = (\bar{A}_{it} L_{it}^{\alpha_1} L_{it-1}^{\alpha_2}) L_{it},
    \]
    \[
    \equiv A_{it},
    \]
    where \(\alpha_1\) and \(\alpha_2\) govern the strength of contemporaneous and historical productivity spillovers.

- **Consumption**
  - Adults have CES preferences over differentiated varieties with EoS \(\sigma\), earn wage \(w_{it}\), have price index \(P_{it}\). Welfare:
    \[
    W_{it} = (\bar{u}_{it} L_{it}^{\beta_1} L_{it-1}^{\beta_2}) \frac{w_{it}}{P_{it}},
    \]
    \[
    \equiv u_{it},
    \]
    where \(\beta_1\) and \(\beta_2\) govern the strength of contemporaneous and historical of amenity spillovers.

- Paper discusses isomorphisms to existing models in the literature
Gravity

- Armington + consumer maximization yields gravity equation for trade:

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

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

- Frechet + welfare maximization yield gravity equation for migration (like commuting in Steve Redding’s lecture):

\[ L_{ijt} = \mu_{ijt}^{-\theta} \prod_{it}^{-\theta} L_{it-1} W_{jt}^{\theta}, \]

- \( \prod_{it} \equiv \left( \sum_{k=1}^{N} \left( W_{kt} / \mu_{ik} \right) \right)^{\frac{1}{\theta}} \) 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 $A(\alpha_1, \beta_1) \equiv$
  \[
  \begin{pmatrix}
    \theta(\alpha_1\sigma+\beta_1(\sigma-1)+1)-(\sigma-1) & \tilde{\sigma}((\sigma-1)(1-(\sigma-1)\alpha_1-\beta_1)+\sigma(\alpha_1\sigma+\beta_1(\sigma-1)+1)) \\
    \sigma+\theta(1-(\sigma-1)\alpha_1-\beta_1) & \sigma+\theta(1-(\sigma-1)\alpha_1-\beta_1) \\
    \theta & \theta(1-(\sigma-1)\alpha_1-\beta_1) \\
    \tilde{\sigma}(\sigma+\theta(1-(\sigma-1)\alpha_1-\beta_1)) & \sigma+\theta(1-(\sigma-1)\alpha_1-\beta_1)
  \end{pmatrix}
  \]

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

- This will occur as long as $\alpha_1$ and $\beta_1$ are sufficiently small.

- Note: Result does not depend on values of $\alpha_2$ and $\beta_2$ (since current generation takes $L_{it-1}$ as given).
Uniqueness if \textit{contemporaneous} spillovers are net dispersive.
How does history shape the distribution of economic activity?

Equilibrium population at any time period in any location:

\[ \gamma \ln L_{it} = C + \sigma \ln \bar{u}_{it} + (\sigma - 1) \ln \bar{A}_{it} - (2\sigma - 1) \ln P_{it} + \sigma \ln \Lambda_{it} + (\alpha_2 (\sigma - 1) + \beta_2 \sigma) \ln L_{i,t-1} \]

where \( \gamma \equiv 1 + \frac{\sigma}{\theta} - (\alpha_1 (\sigma - 1) + \beta_1 \sigma). \)

More people live in locations (if \( \gamma > 0 \)) that have:

- Better (exogenous) geography (\( \bar{A}_{it} \) and \( \bar{u}_{it} \))
- Better (endogenous) market access (for goods \( P_{it} \) and people \( \Lambda_{it} \))
- More people last period (if \( \alpha_2 (\sigma - 1) + \beta_2 \sigma > 0 \))
- Elasticity governed by \( \gamma \) (larger \( \alpha_1 (\sigma - 1) + \beta_1 \sigma \) \( \Rightarrow \) more elastic)
How does history shape the distribution of economic activity?

- From last slide:

\[
\gamma \ln L_{it} = C + \sigma \ln \bar{u}_{it} + (\sigma - 1) \ln \bar{A}_{it} - (2\sigma - 1) \ln P_{it} \\
+ \sigma \ln \Lambda_{it} + (\alpha_2 (\sigma - 1) + \beta_2 \sigma) \ln L_{i,t-1}
\]

- Therefore history persists through three channels:
  - \(L_{i,t-1}\) directly affects \(L_{i,t}\)
  - \(\{L_{j,t-1}\}\) directly affects \(L_{i,t}\) through \(\Lambda_{it}\)
  - \(\{L_{i,t-1}\}\) indirectly effects \(L_{i,t}\) through \(P_{i,t}\)

- Different initial conditions can result in different evolution of the economy even with same geography.
How persistent are historical shocks?

- Define $\mu_{L,t} \equiv \frac{\max_i L_{i,t}/L_{i,t-1}}{\min_i L_{i,t}/L_{i,t-1}}$ for endogenous variable $L$ (similarly $\mu_{W,t}, \mu_{\Pi,t}$).

- Persistence: How are $\mu_t's$ constrained by $\mu_{t-1}'s$?

- **Proposition 2**: Suppose that $\rho(A(\alpha_1, \beta_1)) < 1$. Then:

$$
\begin{pmatrix}
\ln \mu_{L,t} \\
\ln \mu_{W,t} \\
\ln \mu_{\Pi,t}
\end{pmatrix}
\leq
|B^{-1}| \left( I - \tilde{A}(\alpha_1, \beta_1) \right)^{-1} C |B|
\begin{pmatrix}
\ln \mu_{L,t-1} \\
\ln \mu_{W,t-1} \\
\ln \mu_{\Pi,t-1}
\end{pmatrix}
$$

- Note: As $\rho(A(\alpha_1, \beta_1)) \uparrow 1$, the spectral radius of $|B^{-1}| \left( I - \tilde{A}(\alpha_1, \beta_1) \right)^{-1} C |B|$ tends to infinity. So the bound on persistence of shocks grows.
Persistence is potentially \textit{unbounded} as dynamics approach non-uniqueness.

Plot of \( |B^{-1}| \left( I - \tilde{A}(\alpha_1, \beta_1) \right)^{-1} C |B| \) at standard values of \( \sigma \) and \( \theta \)
When can path dependence arise?

- **Recall from Proposition 1:** For any initial population \( \{L_{i0}\} \) there exists a unique (strictly positive) equilibrium if \( \rho(A(\alpha_1, \beta_1)) \leq 1 \).

- **Proposition 3(a):** There exists a unique steady-state if \( \rho(A(\alpha_1 + \alpha_2, \beta_1 + \beta_2)) \leq 1 \).

- **Proposition 3(b):** If \( \rho(A(\alpha_1 + \alpha_2, \beta_1 + \beta_2)) > 1 \), many geographies will have multiple steady-states.

**Implication:** with low \((\alpha_1, \beta_1)\) and high \((\alpha_2, \beta_2)\), can have unique dynamics but multiple steady states.

⇒ “Well-behaved path dependence”.
Suppose $\rho (A (\alpha_1, \beta_1)) \leq 1$ but $\rho (A (\alpha_1 + \alpha_2, \beta_1 + \beta_2)) > 1$.

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 $\alpha_2$...
Phase diagram: 3 symmetric locations, raising $\alpha_2$
Phase diagram: 3 symmetric locations, raising $\alpha_2$
Phase diagram: 3 symmetric locations, raising $\alpha_2$

Location 1: $\alpha_1=0, \alpha_2=0.2, \beta_1=0, \beta_2=0$

Location 2

Location 3

Symmetric locations
Phase diagram: 3 symmetric locations, raising $\alpha_2$
Phase diagram: 3 symmetric locations, raising $\alpha_2$

$\alpha_1 = 0, \alpha_2 = 0.4, \beta_1 = 0, \beta_2 = 0$
Phase diagram: 3 symmetric locations, raising $\alpha_2$

Location 1: $\alpha_1 = 0, \alpha_2 = 0.5, \beta_1 = 0, \beta_2 = 0$

Location 2

Location 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, raising $\alpha_2$

Asymmetric locations (Location 1 has higher amenity)

Location 1
1 = 0, 2 = 0, 1 = 0, 2 = 0

Location 2

Location 3

Location 1
$\alpha_1 = 0, \alpha_2 = 0, \beta_1 = 0, \beta_2 = 0$
Phase diagram: 3 asymmetric locations, raising $\alpha_2$

Asymmetric locations (Location 1 has higher amenity)

Location 1: $\alpha_1 = 0$, $\alpha_2 = 0.1$, $\beta_1 = 0$, $\beta_2 = 0$

Location 2

Location 3
Phase diagram: 3 asymmetric locations, raising $\alpha_2$

Asymmetric locations (Location 1 has higher amenity)

Location 1
$1=0, 2=0.2, 1=0, 2=0$

Location 2

Location 3

$\alpha_1=0, \alpha_2=0.2, \beta_1=0, \beta_2=0$
Phase diagram: 3 asymmetric locations, raising $\alpha_2$
Phase diagram: 3 asymmetric locations, raising $\alpha_2$
Phase diagram: 3 asymmetric locations, raising $\alpha_2$

Asymmetric locations (Location 1 has higher amenity)

$\alpha_1 = 0, \alpha_2 = 0.5, \beta_1 = 0, \beta_2 = 0$
In first example, the 3 stable steady states had identical welfare implications.

But similar intuition holds when the steady states are associated with different welfare levels.

In second example, 3 *asymmetric* locations.

*Implication*: Initial population could cause world to converge to “bad” steady state...

...but note how the “good” steady state has larger basin of attraction.
Ranking steady states

In any steady state, the following notion of welfare is equalized across all locations:

\[ \Omega \equiv E \left[ \max_i (W_i \Pi_i \varepsilon_i) \right] = W_i \Pi_i L_i^{-\frac{1}{\theta}}, \quad \forall i \in \{1, \ldots, N\} \]

Steady state \( B \) is inferior to steady state \( A \) iff

\[ \Omega(\{L_{i0}\}^B) < \Omega(\{L_{i0}\}^A) \]

How does geography affect steady state welfare?
Geography and the welfare costs of path dependence

Let $M \equiv \{ \mu_{ij} \}$ and $T \equiv \{ \tau_{ij}^{1-\sigma} \}$

**Proposition 4:** If $\chi \equiv \alpha_1 + \alpha_2 + \beta_1 + \beta_2 - \frac{1}{\theta} \in \left( 0, \frac{1}{\sigma-1} \right)$ then:

$$\Omega \leq \Omega(\{L_{i0}\}) \leq \bar{\Omega},$$

where:

$$\bar{\Omega} \equiv c_1 \bar{\lambda}_M^\theta \bar{\lambda}_T^{\sigma-1} \| \bar{A}\bar{u} \| \left( \frac{1}{\sigma-1} - \rho \right)^{-1} \bar{L}^\rho \left( \frac{\bar{L}}{N} \right)^{-\frac{1}{\theta}} N^{\frac{1}{2}} \left( \frac{1}{\sigma-1} - \frac{1}{\theta} \right)$$

$$\Omega \equiv c_2 \bar{\lambda}_M^\theta \bar{\lambda}_T^{\sigma-1} \| \bar{A}\bar{u} \| \left( \frac{1}{\sigma-1} - \rho + \frac{1}{\theta} \right)^{-1} \left( \frac{\bar{L}}{N} \right)^{\rho - \frac{1}{\theta}} N^{-\frac{1}{2}} \left( \frac{1}{\sigma-1} - \frac{1}{\theta} \right)$$

and where $\bar{\lambda}_B$ and $\underline{\lambda}_B$ are the largest and smallest eigenvalues of matrix $B$, $\| \bar{A}\bar{u} \|_a \equiv \left( \sum_i (\bar{A}_i \bar{u}_i)^a \right)^{\frac{1}{a}}$ for scalar $a$ and $c_1, c_2$ are constants capturing the dispersion of welfare, population across locations (i.e. $c_1 = 1$ if $W_i = W$).
Ratio of bounds in Proposition 4 place bound on importance of history:

\[
\frac{\Omega}{\tilde{\Omega}} = \frac{c_1}{c_2} \kappa(M)^{1/\theta} \kappa(T)^{1/\sigma - 1} \frac{\|\bar{A}\bar{u}\|}{\|\bar{A}\bar{u}\|} \left(\frac{1}{\sigma - 1} - \rho\right)^{-1} \eta \left(\frac{1}{\sigma - 1} + \rho - \frac{1}{\theta}\right)^{-1},
\]

where:
- \(\kappa(B)\) is the condition number (ratio of abs. val. of max to min eigenvalues) of matrix \(B\).

**Implications:**
- \(\kappa(B)\) measures “sensitivity” of \(Bx = c\) to changes in \(c\) \(\implies\) More sensitive geography of trade costs / migration costs, greater potential importance of history.

\[
\lim_{\rho \searrow \frac{1}{\sigma - 1}} \|\bar{A}\bar{u}\| \left(\frac{1}{\sigma - 1} - \rho\right)^{-1} = \max_i \bar{A}_i \bar{u}_i;
\]

\[
\lim_{\rho \searrow \frac{1}{\tau - 1}} \|\bar{A}\bar{u}\| \left(\frac{1}{\sigma - 1} - \rho + \frac{1}{\theta}\right)^{-1} = \min_i \bar{A}_i \bar{u}_i \implies
\]

Upper (lower) bound places more weight on productivities, amenities in better (worse) locations.
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 only (Commodity Flow Survey)
Estimation

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

- 3-step procedure:
  - Step #1: Assume that trade/migration costs \(\{\tau_{ijt}, \mu_{ijt}\}\) are a function of distance. Estimate those functions via standard gravity equations (both trade and migration).
  - Step #2: recover four key locational characteristics via model inversion
  - Step #3: linear estimation of labor “supply” and “demand” curves to recover \((\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta)\), where residuals are \(\{\bar{A}_{it}, \bar{u}_{it}\}\):
Estimation Step #1: Recovering trade and migration costs

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

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

Table: The gravity distance elasticity for trade and migration

<table>
<thead>
<tr>
<th></th>
<th>Trade (1) 1997</th>
<th>(2) 1850</th>
<th>(3) 1900</th>
<th>Migration (4) 1950</th>
<th>(5) 2000</th>
<th>(6) Pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log distance</td>
<td>-1.353***</td>
<td>-2.157***</td>
<td>-1.798***</td>
<td>-1.890***</td>
<td>-1.511***</td>
<td>-1.755***</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.080)</td>
<td>(0.035)</td>
<td>(0.025)</td>
<td>(0.017)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Origin-year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Destination-year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.901</td>
<td>0.689</td>
<td>0.793</td>
<td>0.857</td>
<td>0.908</td>
<td>0.864</td>
</tr>
<tr>
<td>Observations</td>
<td>2091</td>
<td>626</td>
<td>1991</td>
<td>2152</td>
<td>2304</td>
<td>7073</td>
</tr>
</tbody>
</table>
Estimation Step #2: Model inversion

- Define $T_{ij} \equiv \frac{1}{\tau_{ij}^{1-\sigma}} = \text{dist}_{ij}^{-(\sigma-1)\kappa}$, $M_{ij} \equiv \mu_{ij}^{1-\theta} = \text{dist}_{ij}^{-\theta \lambda_t}$

- Equilibrium conditions (with $p_{it} \equiv \frac{w_{it}}{A_{it}}$ and $Y_{it} \equiv w_{it}L_{it}$):

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

  \[
  P_{jt}^{\sigma-1} = \sum_i T_{ij} \left( p_{it}^{\sigma-1} \right)^{-1}
  \]

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

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

- Proposition 4: Given observed $\{Y_{it}, L_{it}, L_{it}^{-1}, T_{ij}, M_{ij}\}$, there exists unique (to-scale) endogenous characteristics $\{p_{it}^{\sigma-1}, P_{it}^{\sigma-1}, W_{it}^{\theta}, \Pi_{it}^{\theta}\}$. 
Estimation Step #3: The estimating equation

- Following log-linear labor supply and demand system:

\[
\ln Y_{it} = \left( \frac{\sigma - 1}{\sigma} \right) (1 + \alpha_1) \ln L_{it} + \left( \frac{\sigma - 1}{\sigma} \right) \alpha_2 \ln L_{it, \text{lag}} + \frac{1}{\sigma} \ln P_{it}^{1-\sigma} + \frac{\sigma - 1}{\sigma} \ln A_{it}
\]

\[
\ln Y_{it} = (1 - \beta_1) \ln L_{it} + (-\beta_2) \ln L_{it, \text{lag}} + \frac{1}{\theta} \ln W_{it}^{\theta} + \left( \frac{1}{1 - \sigma} \right) \ln P_{it}^{1-\sigma} - \ln \bar{u}_{it}
\]

- Notes:
  
  - Recovered “market access” parameters \( P_{it}^{1-\sigma} , W_{it}^{\theta} \) affect intercept (not slope).
  
  - \( \bar{A}_{it} \) shifts demand (but not supply); \( \bar{u}_{it} \) shifts supply (but not demand).
  
  - 2SLS can identify elasticities \( \alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta \) and fundamentals \( \{ \bar{A}_{it}, \bar{u}_{it} \} \), despite potentially severe issue of multiplicity.
  
  - Whole thing is an extended version of Rosen-Roback framework discussed in Ed Glaeser’s lecture (set \( \sigma = \theta \to \infty \) to recover that case)
Estimation Step #3: Instruments

- **IV for demand curve**: Changes in amenities
  - 2nd order polynomial of: average maximum (minimum) temperature in hottest (coldest) month.
  - *Relevance*: Over time, technological advances have mitigated dis-amenity of extreme temperatures.
  - *Exclusion restriction*: Extreme temperature do not affect productivities (cond’l on controls, fixed effects).

- **IV for supply curve**: Changes in productivity
  - Mean and Std.dev. of difference in potential yield of high vs low yield corn, high yield soy vs low yield wheat.
  - *Relevance*: technological advances in agriculture, secular changes in crop mix affects productivity.
  - *Exclusion restriction*: Relative local potential yields do not affect amenities (cond’l on controls, fixed effects).
### Estimation results: Productivity Spillovers

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimated parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Productivity spillover $\alpha_1$</td>
<td>1.149</td>
<td>0.204***</td>
<td>0.104***</td>
</tr>
<tr>
<td></td>
<td>(3.156)</td>
<td>(0.042)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>Productivity spillover $\alpha_2$</td>
<td>0.142</td>
<td>0.070*</td>
<td>0.062*</td>
</tr>
<tr>
<td></td>
<td>(0.272)</td>
<td>(0.040)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>Elasticity of substitution $\sigma$</td>
<td>1.700</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.930)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Assumed parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elasticity of substitution $\sigma$</td>
<td></td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td><strong>Fixed effects:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-county</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Region-year</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Min SW 1st-stage F-stat</td>
<td>24.258</td>
<td>66.520</td>
<td>66.520</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.628</td>
<td>0.081</td>
<td>0.036</td>
</tr>
<tr>
<td>Observations</td>
<td>15640</td>
<td>15640</td>
<td>15640</td>
</tr>
</tbody>
</table>

- NB: A&D treat estimates in column (2) as baseline
## Estimation results: Amenity Spillovers

### Estimated parameters:

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amenity spillover $\beta_1$</td>
<td>0.794 (3.508)</td>
<td>1.432 (1.160)</td>
<td>-0.120 (0.402)</td>
<td>-0.372 (0.396)</td>
<td>-0.498 (0.394)</td>
<td>1.378 (1.152)</td>
<td>-0.115 (0.403)</td>
<td>-0.366 (0.397)</td>
<td>-0.492 (0.395)</td>
</tr>
<tr>
<td>Amenity spillover $\beta_2$</td>
<td>0.483 (0.664)</td>
<td>0.346 (0.288)</td>
<td>0.446* (0.262)</td>
<td>0.462* (0.257)</td>
<td>0.470* (0.255)</td>
<td>0.358 (0.286)</td>
<td>0.454* (0.262)</td>
<td>0.470* (0.258)</td>
<td>0.478* (0.256)</td>
</tr>
<tr>
<td>Migration elasticity $\theta$</td>
<td>0.744 (2.181)</td>
<td>0.490* (0.284)</td>
<td>0.505* (0.300)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elasticity of substitution $\sigma$</td>
<td>0.175 (5.085)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Assumed parameters:

<table>
<thead>
<tr>
<th></th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>9</th>
<th>9</th>
<th>9</th>
<th>9</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity of substitution $\sigma$</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Migration elasticity $\theta$</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

### Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-county</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Region-year</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Min SW 1st-stage F-stat</td>
<td>0.912</td>
<td>5.636</td>
<td>5.043</td>
<td>5.043</td>
<td>5.043</td>
<td>5.043</td>
<td>5.043</td>
<td>5.043</td>
<td>5.043</td>
<td>5.043</td>
</tr>
<tr>
<td>R-squared</td>
<td>-0.800</td>
<td>-4.473</td>
<td>-1.151</td>
<td>-1.454</td>
<td>-1.468</td>
<td>-4.467</td>
<td>-1.180</td>
<td>-1.504</td>
<td>-1.528</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>15640</td>
<td>15640</td>
<td>15640</td>
<td>15640</td>
<td>15640</td>
<td>15640</td>
<td>15640</td>
<td>15640</td>
<td>15640</td>
<td></td>
</tr>
</tbody>
</table>

- NB: A&D treat estimates in column (4) as baseline
Baseline estimates consistent with well-behaved potential path dependence and substantial persistence

- Recall Propositions 1-3
History explains most of variation in where people live

\[
\gamma \ln L_{it} = C + \sigma \ln \bar{u}_{it} + (\sigma - 1) \ln \bar{A}_{it} - (2\sigma - 1) \ln P_{it} + \sigma \ln \Lambda_{it} + (\alpha_2 (\sigma - 1) + \beta_2 \sigma) \ln L_{i,t-1}
\]

\[\text{exogenous (local) geography} \quad \text{endogenous (global) geography}\]

\[\text{direct impact of history}\]

**Figure:** Variance decomposition of observed population distribution
Constructing counterfactual histories

- Desire now to quantify the role of “good history” in benefiting long-run success of particular locations in US
- Pair each location with a “buddy” that has closest population in 1900.
  - E.g. Milwaukee (1900 pop: 285,315) and New Orleans (1900 pop: 287,104) are “buddies”.
- For each counterfactual history, randomize (without replacement) $\bar{A}_{i,1900}$ between buddies.
  - E.g. $\bar{A}_{Mil,1900} = 0.42$, $\bar{A}_{NO,1900} = 0.20$, so $\approx 50\%$ of time, Milwaukee gets unlucky N.O. productivity (and vice versa).
- Hold constant $\bar{A}_{it}$ for all other years and $\bar{u}_{it}$ for all years.
  - Implication: Differences in economic activity in year 2000 driven only through history (not geography).
- Intuition: Small perturbations construct plausible counterfactual 1900 population distributions.
Alternative histories: Examples, 1850

(a) Actual history

(b) Alternative history #1

(c) Alternative history #2

(d) Alternative history #3
Alternative histories: Examples, 1900

(e) Actual history
(f) Alternative history #1

(g) Alternative history #2
(h) Alternative history #3
Alternative histories: Examples, 1950

(i) Actual history

(j) Alternative history #1

(k) Alternative history #2

(l) Alternative history #3
Alternative histories: Examples, 2000

(m) Actual history

(n) Alternative history #1

(o) Alternative history #2

(p) Alternative history #3
Takeaway 1: Historical shocks are very persistent

- 1850
- 1900
- 1950
- 2000
Takeaway 1: Historical shocks are very persistent

**Figure:** Correlation between observed and counterfactual populations
Takeaway 2(a): Impact of history was heterogeneous

**Figure:** How sensitive was each location to a historical shock?

Notes: The y-axis is the estimated elasticity of the population in the year 2000 to the productivity shock in 1900, i.e. \( \frac{\partial \ln L_{i,2000}}{\partial \ln A_{i,1900}} \), estimated across the 200 counterfactual histories. The x-axis is the actual (log) population in the year 2000. Quadratic fit, weighting observations by the inverse of the variance of the estimated elasticity.
Takeaway 2: Impact of history was heterogeneous

Figure: How lucky was each location?

Notes: Fraction of counterfactual histories where population rank was lower than actual history. Red (blue) indicates actual population rank was higher (lower) than most counterfactual histories.
Take away 3: Population more sensitive than welfare

**Figure**: Variation in population and welfare across counterfactual histories

(a) 1900  
(b) 1950  
(c) 2000

*Notes*: Scatter plot of standard deviation of log welfare and log population across counterfactual histories for each location over time.
Take away 3: Population more sensitive than welfare

Figure: Welfare versus population sensitivity to historical shocks

Notes: Kernel density plots across locations of elasticity of population and welfare in 2000 to 1900 shocks, i.e. $\frac{\partial \ln L_i,2000}{\partial \ln A_i,1900}$ and $\frac{\partial \ln W_i,2000}{\partial \ln A_i,1900}$.
Stepping back... does history matter in this model?

- Yes: Small historical shocks have large and persistent impacts on where people live for hundreds of years.
  - ... but impacts on welfare much smaller than impacts of where people live.

- No: No evidence of path dependence from shocks Allen and Donaldson (2020) consider.
  - ... but remains theoretically possible from other shocks (and bounds suggest very large welfare consequences can’t be ruled out).
Lots of great textbooks cover the themes of this lecture:

- Fujita, Krugman and Venables (1999)
- Fujita and Thisse (2002)
- Brakman, Garretsen, and van Marrewijk (2003)
- Glaeser (2008)
- Combes, Mayer and Thisse (2009)

Elsevier Handbooks of Regional and Urban Economics (especially: Henderson and Thisse eds. (2004), and Duranton, Henderson and Strange eds. (2015))