INTERNATIONAL TRADE - ECON 245

FABIAN ECKERT

MIGRATION

ALLEN DONALDSON

SETUP

- Armington trade model with migration
- Contribution 1: Model is dynamic, so object are indexed by time period t
  - Agents live for two periods: childhood and adulthood
    - Overlapping generations-like model but without altruism
    - Kills dynamics for individual decisions, but allows to think about time
- Contribution 2: productivities and amenities depend not only on current but also past local populations; potential for different paths of the economy!

MIGRATION

- Consider a distribution of workers across locations L_{rt} at time t
- Each of these workers has an "invisible" child during time period t
- At the end of time t the parents die and the children choose a location to work
- Migration costs μ_{ij} to move from i to j, paid in utility terms, wlog $\mu_{ii} = 1, \mu_{ij} \geq 1 \forall j \neq i$
- Each agent ω has idiosyncratic preferences for each destination η_j^ω
  - Total utility in each destination is give by: $W_j^\omega = U_j \frac{w_j}{P_j} \eta_j^\omega$

MIGRATION

- Each agent draws the η_j^ω i.i.d. from a Frechet distribution with scale parameter 1 ("mean") and shape parameter θ ("dispersion")
- An individual young adult in birthplace location i hence solves:

$$j^\star = \arg\max\{W_j^\omega\}$$

- But then we obtain an analytic expression for the fraction of young adults choosing destination j:

$$\pi_{ijt} = \frac{\left(\mu_{ij}U_j w_j/P_j\right)^\theta}{\sum_k \left(\mu_{ik}U_k w_k/P_k\right)^\theta} \equiv \frac{\left(\mu_{ij}U_j w_j/P_j\right)^\theta}{\Lambda_i}$$

EQUILIBRIUM

- The equilibrium system is somewhat simpler than in the free mobility Armington case
- Labor/goods market clearing:

$$L_i w_i = \sum_j \lambda_{ij} L_j w_j$$

  - Where λ_{ij} are the standard Armington trade shares
- The spatial equilibrium condition is now simply

$$L_{i,t} = \sum_k \pi_{k,i} L_{k,t-1}$$

  - Where the migration shares $\pi_{k,i}$ summarize optimal migration choices of agents

OBSERVATIONS

- Since $\mu_{ii} = 1$ and $\mu_{ij} \geq 1 \forall j \neq i$ there is an incentive to forego wage differences and just stay where you were born
- The fraction of young adults moving is also the ex-ante probability of an individual worker moving before they learn their productivity
- The fraction of workers moving to k relative to those moving to k' is unaffected by the fraction moving to k'', i.e., $\pi_{ik}/\pi_{ik'}$ is independent of $\pi_{ik''}$. Realistic?
- Next two observations deserve their own slides…

MIGRATION GRAVITY

- We now get a true gravity equation for migration flows!
- The *number* of people moving from i to j in period t to work there can be written as:

$$L_{ij,t} = \frac{\left(\mu_{ij} U_j w_j / P_j\right)^{\theta}}{\Lambda_i} L_{i,t-1}$$

- Taking logs: $\log(L_{ij,t}/L_{i,t-1}) = \theta \log(\mu_{ij}) + \theta \log(U_j w_j / P_j) + \log(1/\Lambda_i)$
- Parameterising migration costs as a function of distance: $\mu_{ij} = d_{ij}^{\kappa}$

$$\log(\pi_{ij,t}) = \theta\kappa \log(d_{ij}) + \varphi_j + \xi_i + \epsilon_{ij}$$

MIGRATION GRAVITY AND THE DATA

- This is usually estimated using a Poisson Maximum Likelihood estimator (see Silva and Tenreyro 2006, paper+stata package!)
  - Rewrite the equation: $\pi_{ij,t} = \exp(\theta\kappa \log(d_{ij}) + \varphi_j + \xi_i)\tilde{\epsilon}_{ij}$
    - This looks a lot like Poisson probability mass function!
    - Easy to construct likelihood of observed migration behavior
- Running these regression in the data:
  - The US census provides state of birth and current state for workers for the last 200 years – can easily construct these lifetime migration probabilities!
  - Linked historical census data another great way to run these regressions

GROSS VERSUS NET MIGRATION FLOWS

- An essential distinction in models of migration is net versus gross flows
- Gross flows between locations i and j are given by:

$$L_{ij,t} = \pi_{ijt} L_{i,t-1} \quad \text{and} \quad L_{ji,t} = \pi_{jit} L_{j,t-1}$$

  - So there are flows in both directions!
- Net flows are then given by:

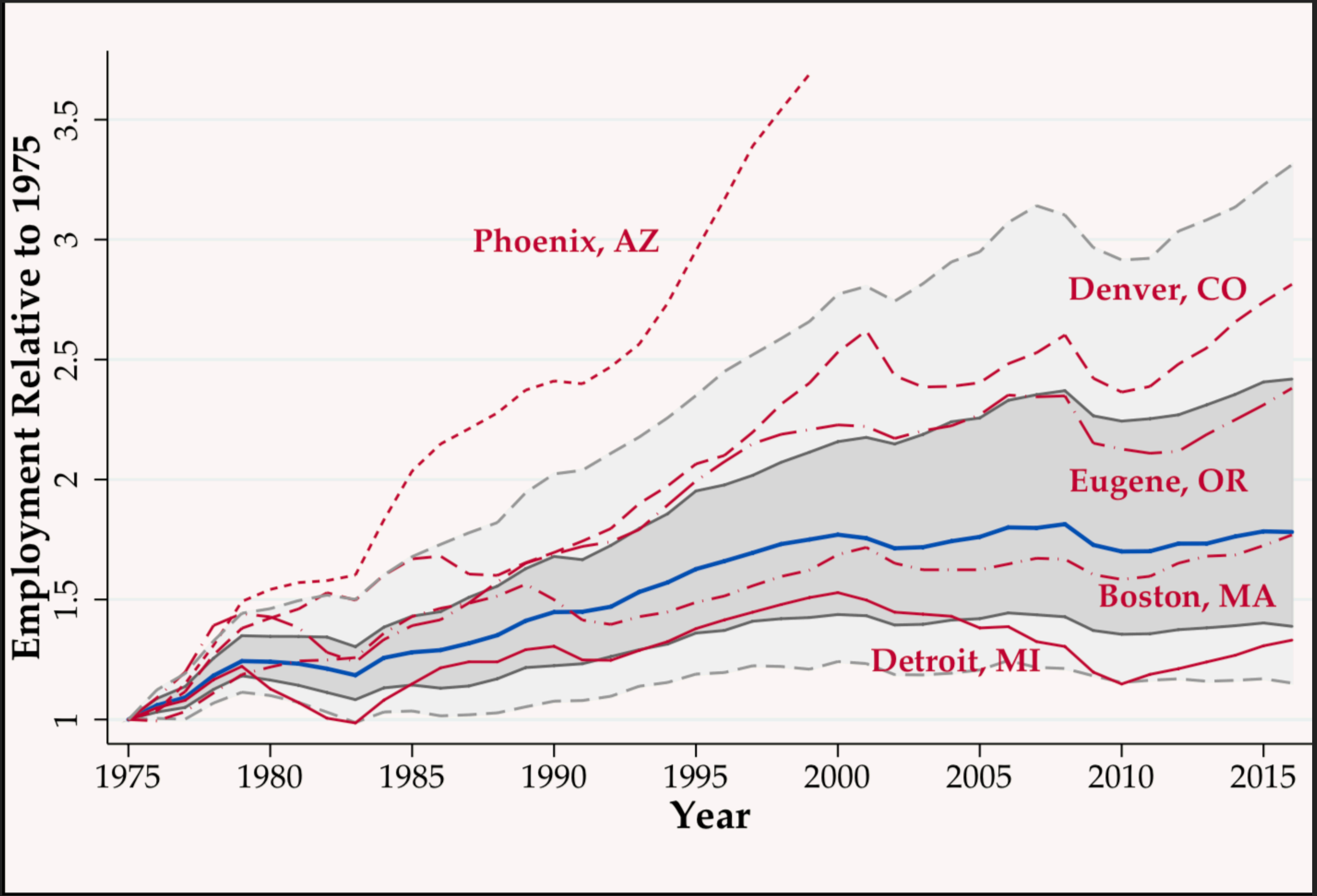
$$\Delta = L_{ij,t} - L_{ji,t} = \pi_{ijt} L_{i,t-1} - \pi_{jit} L_{j,t-1}$$

  - Net flows are only positive for one of the two locations

GROSS VERSUS NET MIGRATION FLOWS

- Suppose we are in a long run steady state of the model
  - Then net flows are zero between regions but there are still gross flows.
- Now suppose we raise A_i for some location i
  - This sets into motion *several* periods of adjustment in which net flows between regions are non-zero
    - In particular there will be positive inflows into region i from the other regions

GRAPH FROM WALSH (2020): NET POPULATION GAINS ACROSS US CITIES



ADDITIONAL BELLS AND WHISTLES: SLOW DOWN ADJUSTMENT DYNAMICS

- We can add a “fixed cost of moving”: suppose each period only a fraction $\lambda \in (0,1)$ make a moving decisions
  - Then the total outflow of region i is given by:

$$L_{i,t} = \sum_k \pi_{k,i} \lambda L_k + (1 - \lambda) L_{i,t-1}$$

  - This way can get arbitrarily slow adjustment dynamics even without forward looking agents!

PATH DEPENDENCE IN ALLEN AND DONALDSON

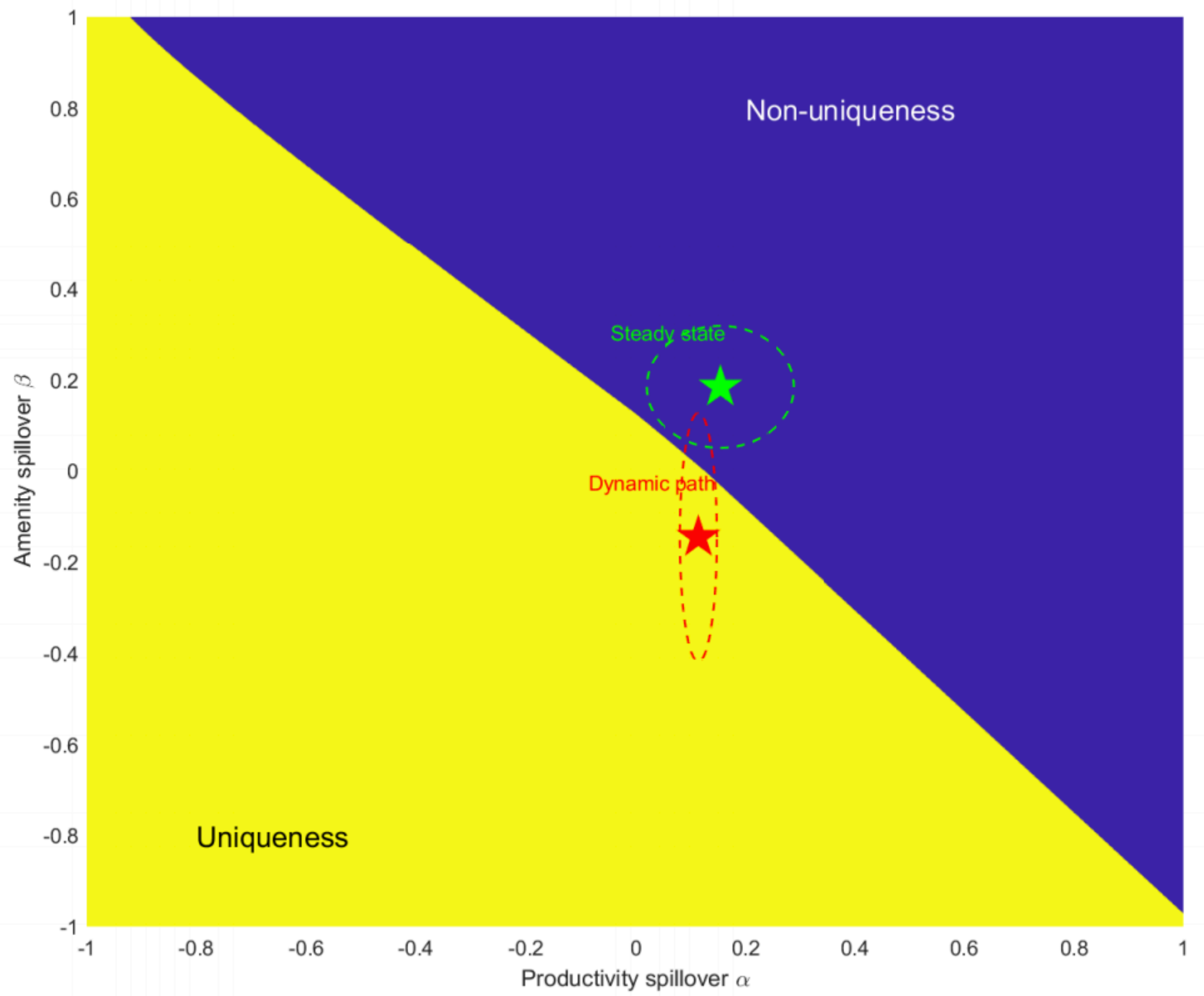
- One you have set up this overlapping generations structure, can think seriously about *path dependence*
- AA introduce path dependence in amenities and productivities as follows:

$$A_{it} = \bar{A}_{it} L_{it}^{\alpha_1} L_{it-1}^{\alpha_2} \quad \text{and} \quad u_{it} = \bar{u}_{it} L_{it}^{\beta_1} L_{it-1}^{\beta_2}$$

- Suddenly initial conditions matter: two different worker distributions at time $t-1$ entail different distributions of productivities and amenities today!
  - Given initial conditions get *unique dynamic path*, but steady state differs by initial condition so there are *multiple steady states*: **history can now matter!**

PATH DEPENDENCE IN ALLEN AND DONALDSON

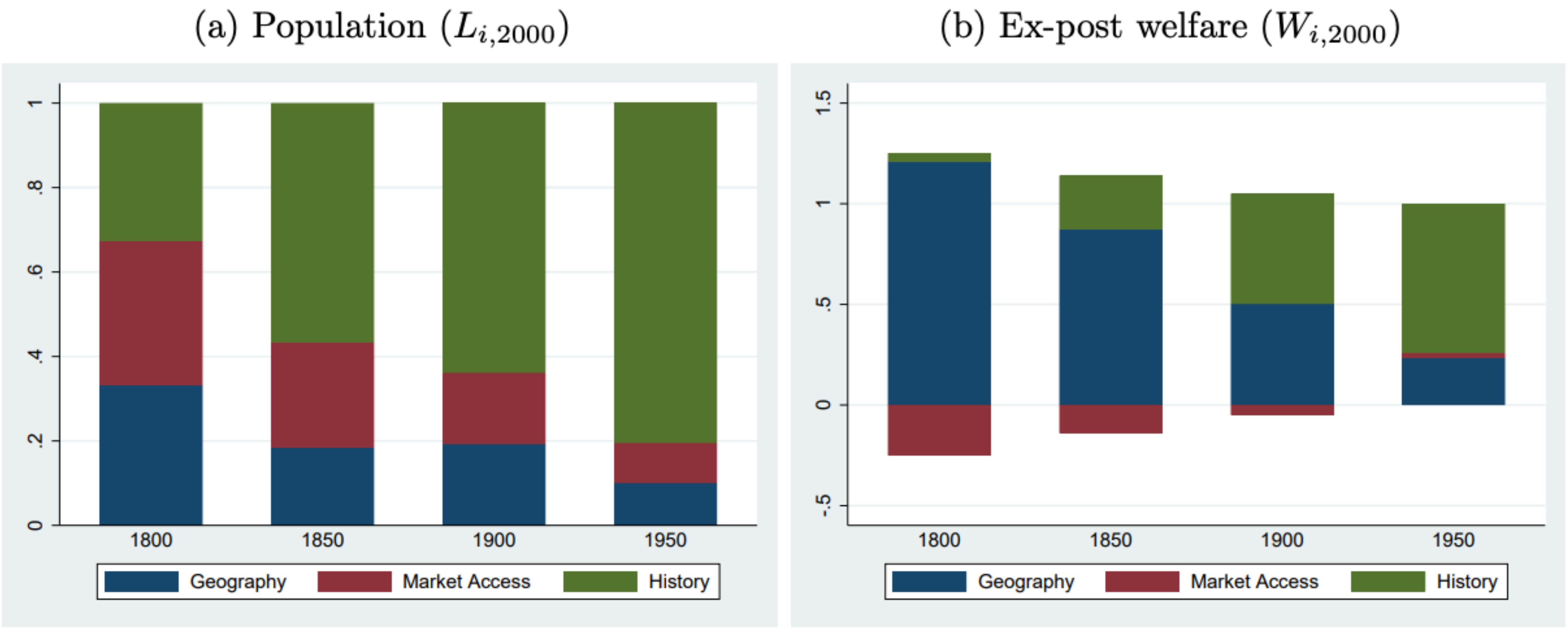
Figure 3: Agglomeration spillover parameter estimates



Notes: This figure illustrates the parameter estimates (holding σ and θ constant) obtained in Section 3.3. The red star denotes $\widehat{\alpha}_1$ and $\widehat{\beta}_1$, which lies in the yellow region of equilibrium uniqueness following Proposition 1. The green star denotes $\widehat{\alpha}_1 + \widehat{\alpha}_2$ and $\widehat{\beta}_1 + \widehat{\beta}_2$, which lies in the blue region, indicating the possibility of multiple steady-states following Proposition 2. Confidence intervals are shown with dashed lines.

PATH DEPENDENCE IN ALLEN AND DONALDSON

Figure 4: How much of the spatial distribution of economic activity today is due to history?



Notes: This figure presents the variance decomposition of the observed spatial distribution of economic activity in the year 2000 into three components, as per equation (33): geography fundamentals (i.e. the complete history of realizations of productivities $\bar{A}_{it}$ and amenities $\bar{u}_{it}$ from $t = 0$ until the present), market access (i.e. the complete history of goods market access P_{it} and labor market access Λ_{it} from $t = 0$ until the present), and history (i.e. the population distribution in $t = 0$, L_{i0}). The decompositions shown correspond to four choices of initial year $t = 0$: 1800, 1850, 1900, and 1950. Panel (a) presents the decomposition for the observed distribution of population in the year 2000 ($L_{i,2000}$), and panel (b) presents the equivalent for ex-post welfare ($W_{i,2000}$).

RELATED: PORTAGE AND PATH DEPENDENCE BY BLEAKLEY AND LIN

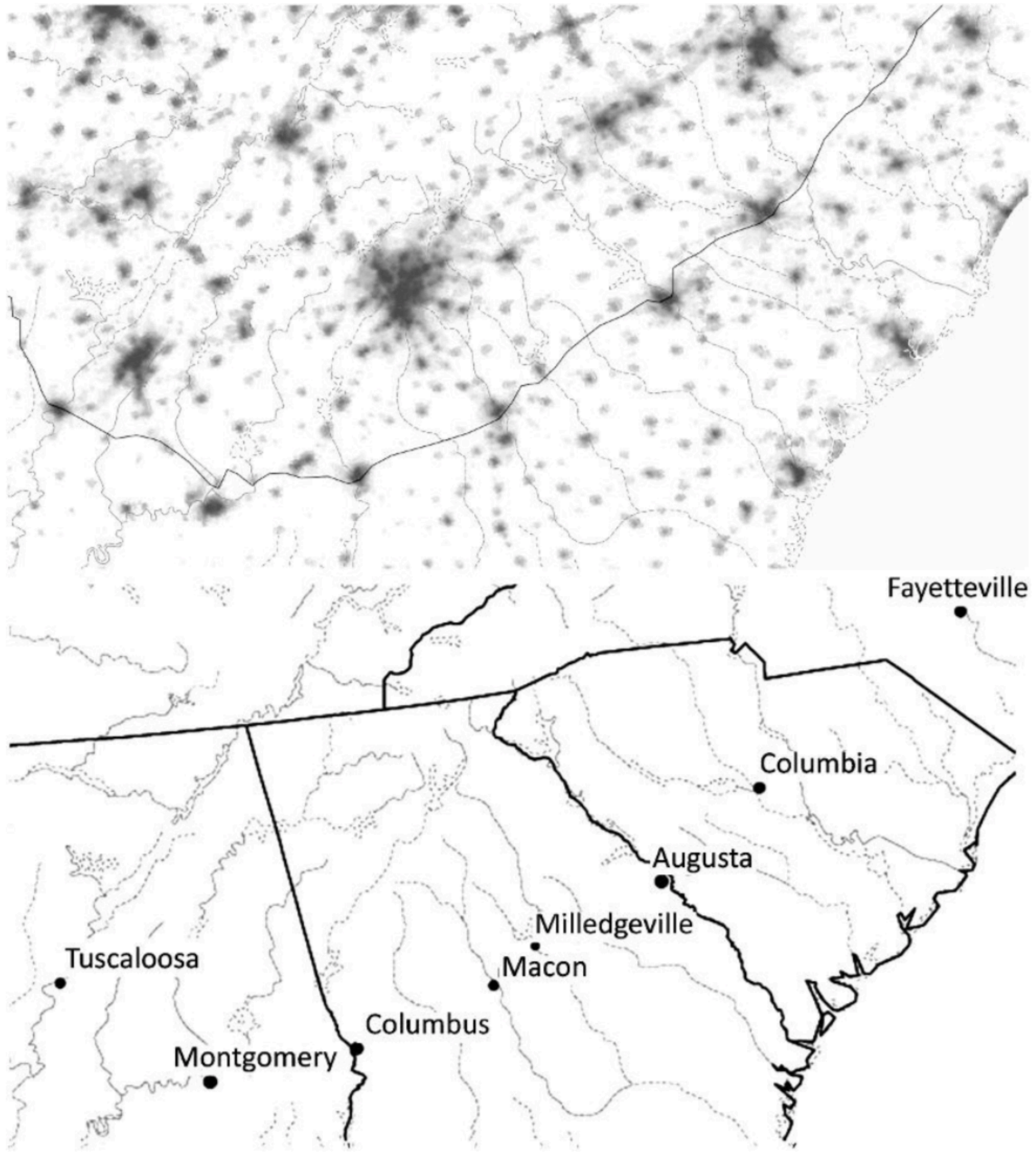


FIGURE II

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.

CALIENDO, DVORKIN, PARRO

INTRODUCTION

- Now we make the migration decision forward-looking, i.e., we have infinitely lived agents.
- The difficulty: agents have to predict the path of wages and rents in each locations to make their moving decisions *today*
  - E.g.: Could be optimal to move to Denver today, because wages are high AND its close to Salt Lake City which will do well in three decades from now
    - There is now an *option value* to being in each location!

SETUP

- Armington trade model with migration; perfect competition; no forward looking decisions by firms
- Novelty is on the worker side:
  - Workers live forever, “dynasties” (like Allen Donaldson but *with* altruism)
  - Discount the future with β
  - Redraw new preference shocks for each location each period
- Idiosyncratic amenity shocks are Gumbel distributed

MIGRATION DECISION

- The utility of worker ω in location i today then looks as follows:

$$v_{i,t}^{\omega} = W_i + \max_j \left(\beta E_{\eta_{t+1}}[v_{j,t+1}^{\omega}] - \mu_{ij} + \theta \eta_{jt}^{\omega} \right) \quad \text{where} \quad W_i = u_i w_i / P_i$$

- Now define $V_{i,t}$ as the expected lifetime value of an agent in i before learning his idiosyncratic preference shock: $V_{i,t} = E_{\eta_t}[v_{i,t}^{\omega}]$

$$V_{i,t} = W_i + E_{\eta_t}[\max_j \left(\beta V_{j,t+1} - \mu_{ij} + \eta_{jt}^{\omega} \right)] = W_i + \theta \log \left[\sum_j \exp \left(\beta V_{j,t+1} - \mu_{ij} \right)^{\frac{1}{\theta}} \right]$$

- So $V_{j,t+1}$ summarizes the *option* value of being in location j tomorrow taking into account all possible future draws of idiosyncratic shocks

MIGRATION DECISION

- The fraction of workers moving from i to j is now like the Gumbel one we discussed, BUT involves the continuation value:

$$\pi_{ij} = \frac{\exp\left(\beta V_{j,t+1} - \mu_{ij}\right)^{\frac{1}{\theta}}}{\sum_j \exp\left(\beta V_{j,t+1} - \mu_{ij}\right)^{\frac{1}{\theta}}}$$

- So now workers assess destinations not just for their current wage, but for the path of future wages and the strategic migration position they offer!
  - "Option value" of each location matters
  - Option value does not depend on idiosyncratic shocks since they a redrawn each period

STEADY STATE EQUILIBRIUM

- Labor/goods market clearing:

$$L_i w_i = \sum_j \lambda_{ij} L_j w_j$$

- The spatial equilibrium condition is now simply

$$L_{j,t} = \sum_k \pi_{k,j} L_{k,t-1}$$

- In steady stage, by definition $V_j = V_{j,t} = V_{j,t-1}$ so can solve for the option value:

$$V_j = W_j + \theta \log \left[\sum_k \exp\left(\beta V_k - \mu_{jk}\right)^{\frac{1}{\theta}} \right]$$

SOLVING FOR THE TRANSITION

- Suppose now some fundamental changes $\Theta = \{A_{it}, u_{it}, \tau_{ijt}, \mu_{ijt}\}$
  - Could be one-off changes or and entire path of changes
- In theory can then solve in two steps:
  - Step 1: Compute new steady state where $V_i = V_{i,T} = V_{i,T-1}$ for some T using previous slide
  - Step 2: Guess path of value functions $\{V_{it}\}_{t=0}^{T}$ and then simulate forward from initial conditions in period 0, iterate

SOLVING FOR THE TRANSITION: DYNAMIC HAT ALGEBRA

- Solving the equilibrium of the model requires knowing Θ at each point in time
- As we added countries, regions, sectors, the number of parameters grows fast
- Rewrite the model in changes: given data for the initial period, can solve for entire transition to new steady state without knowing Θ

SOLVING FOR THE TRANSITION: DYNAMIC HAT ALGEBRA

- The model in changes:

$$\exp(V_{it+1} - V_{it}) = (\hat{w}_{it+1}/\hat{P}_{it+1}) \left[\sum_j \exp\left(V_{jt+2} - V_{jt+1}\right)^{\beta/\theta} \pi_{ijt} \right]$$

- Migration shares:

$$\pi_{ijt+1} = \frac{\pi_{ijt} \exp(V_{jt+2} - V_{jt+1})^{\beta/\theta}}{\sum_k \pi_{ikt} \exp(V_{kt+2} - V_{kt+1})^{\beta/\theta}}$$

- And migration itself:

$$L_{i,t} = \sum_k \pi_{k,i} L_{k,t-1}$$

  - where as always $\hat{x}_{t+1} = x_{t+1}/x_t$

ALGORITHM IN BROAD SKETCHES

- Guess path for $\{V_{it+1} - V_{it}\}_{t=0}^{T}$, T is also a guess. Note that we have data at time $t = 0$ and that $\{\hat{\Theta}_t\}_{t=0}^{T}$ which is the "shock" the response to which we compute.
- Use this to solve for population in each region in each location at each point in time using the migration shares in changes
- Use these populations to solve static labor marker equilibrium at each point in time.
- Use new steady state in T and path of wages and prices to infer a new guess for $\{V_{it+1} - V_{it}\}_{t=0}^{T}$

BRYAN AND MORTEN

INTRODUCTION

- Why we discuss the paper:
  - It produces a gravity equation for migration, while still being a static model
  - It highlights selection on productivity, so far we have only seen selection on idiosyncratic preferences
  - It introduces *correlated* Frechet shocks into the migration literature
  - It highlights the limits of the Frechet approach in modeling selection

EMPIRICAL OBSERVATIONS

- *Fact 1*: The further the distance between origin and destination, the lower the share of origin workers who choose to migrate.
- *Fact 2*: Controlling for origin and destination fixed effects, workers the migrated further distance earn more on average.
- *Fact 3*: Controlling for origin and destination fixed effects, if more people migrate on a route their average wage is lower.
- *Fact 4*: Wages are higher for longer distance migrants due to selection.
- *Fact 5*: Controlling for origin fixed effects, high destination amenities, yields lower average wages for migrants

SETUP

- Economy consists of N locations
- Standard Armington setup: region-specific varieties, CES, perfect competition
- A mass of workers L_o is born into each location o
- Workers receive idiosyncratic productivity shocks for each possible destination
  - Reduced form way of modeling that some workers are productive in some locations depending on their skills and local industrial structure
- Workers then choose their labor market of employment indexed by d

LOCATION CHOICES

- Worker i from origin o *could* supply the following human capital in destination d:

$$h^i_{od} = s^i_d q_o$$

  - Where q_o can be thought of as a measure of the quality of education in origin

- The total utility in destination d of someone born into o is:

$$U^i_{od} = u_d \mu_{od} w_d h^i_{od}$$

  - u_d: amenities, w_d: wages per unit of HC, μ_{od}: moving cost, $\mu_{oo} = 1, \mu_{od} < 1 \, \forall o \neq d$

LOCATION CHOICES

- The idiosyncratic productivity shocks are drawn from a *multivariate* Frechet

$$F(s_1, \ldots, s_N) = \exp\{-\left[\sum_d s_d^{\frac{\theta}{1-\rho}}\right]^{1-\rho}\} \equiv \exp\{-\left[\sum_d s_d^{\theta}\right]^{1-\rho}\}$$

- Observations:
  - The entries of the vector of shocks for an agent are no longer i.i.d!
  - $1/\theta$ measures dispersion of shocks: importance of comparative advantage
  - ρ measures the correlation between shocks of a given worker
    - $\rho \to 1$ talent becomes unidimensional; $\rho \to 0$ back to i.i.d shocks!

LOCATION CHOICES

- The resulting location choices have familiar expressions:

$$\pi_{od} = \frac{(u_d \mu_{od} w_d q_o)^\theta}{\sum_d (u_d \mu_{od} w_d q_o)^\theta} = \frac{(u_d \mu_{od} w_d)^\theta}{\sum_d (u_d \mu_{od} w_d)^\theta}$$

- Importantly, origin effects that are not indexed by destination (q_o) cancel out!
- So different from the models thus far we get a gravity equation for migration flows:

$$L_{od} = \pi_{od} L_o \Rightarrow \log(L_{od}) = \delta^o + \delta^d + \theta \log(\mu_{od})$$

- Can run this exact regression in US Census data from IPUMS (across states)

SELECTION

- The Fréchet assumptions leads to two more nice analytical expressions:

$$E(s_d \mid d^\star = d) = \pi_{od}^{-\frac{1}{\theta}}\bar{\Gamma} \quad \text{and} \quad \bar{w}_{od} = w_d q_o \pi_{od}^{-\frac{1}{\theta}}\bar{\Gamma}$$

  - Selection: the fewer people move the higher their average productivity
- Only workers with very specific talent make long distance moves
  - Model predicts that American expats should earn more on average than Americans at home
- Large wage gap across locations may then not imply high returns to moving people: moving people would lower their average productivity and wages

SELECTION

- One other implications of the model

$$\frac{\bar{w}_{od}}{\bar{w}_{od'}} = \frac{u_d}{u_{d'}} \frac{\mu_{od}}{\mu_{od'}}$$

- Observations:
  - No difference in wages across destinations without frictions/amenities
  - Observed wage gaps then are due to frictions that induce selection
    - Young (2013) spatial wage gaps are all about selection, this is not due here
  - Strange implication: destination wage difference do not show up: inherently productive locations attract the, on average, least productive workers.
    - Assignment models (see Gaubert 2018, or Bilal Rossi-Hansberg 2020) have more realistic notion of sorting)