A Dynamic Spatial Knowledge Economy

Levi Crews (Princeton)

September 2023

- **Micro estimates: Workers learn more in big cities**

  [Glaeser and Maré (2001), Baum-Snow and Pavan (2012), Wang (2016), De la Roca and Puga (2017)]

  > *"[W]e find that workers in **bigger cities** . . . obtain an immediate static [earnings] premium and accumulate **more valuable experience**. The additional value of experience in bigger cities **persists** after leaving and is **stronger** for those with higher initial ability."* (De la Roca and Puga, 2017)

- **Micro estimates: Workers learn more in big cities** $\rightarrow$ **What are the GE effects?**
  [Glaeser and Maré (2001), Baum-Snow and Pavan (2012), Wang (2016), De la Roca and Puga (2017)]

  *"[W]e find that workers in **bigger cities** ... obtain an immediate static [earnings] premium and accumulate **more valuable experience**. The additional value of experience in bigger cities **persists** after leaving and is **stronger** for those with higher initial ability."* (De la Roca and Puga, 2017)

Introduction

- **Micro estimates: Workers learn more in big cities → What are the GE effects?**
  [Glaeser and Maré (2001), Baum-Snow and Pavan (2012), Wang (2016), De la Roca and Puga (2017)]

  > *"[W]e find that workers in **bigger cities** ... obtain an immediate static [earnings] premium and accumulate **more valuable experience**. The additional value of experience in bigger cities **persists** after leaving and is **stronger** for those with higher initial ability."* (De la Roca and Puga, 2017)

- **We think learning drives growth → How does the spatial dist. matter for growth?**
  [Lucas (2009), Lucas and Moll (2014), Buera and Lucas (2018), Gabriel and Lucas (2019)]

Introduction

- **Micro estimates: Workers learn more in big cities → What are the GE effects?**
  [Glaeser and Maré (2001), Baum-Snow and Pavan (2012), Wang (2016), De la Roca and Puga (2017)]

  *"[W]e find that workers in **bigger cities** . . . obtain an immediate static [earnings] premium and accumulate **more valuable experience**. The additional value of experience in bigger cities **persists** after leaving and is **stronger** for those with higher initial ability."* (De la Roca and Puga, 2017)

- **We think learning drives growth → How does the spatial dist. matter for growth?**
  [Lucas (2009), Lucas and Moll (2014), Buera and Lucas (2018), Gabriel and Lucas (2019)]

- **If space matters → Spatial policy must weigh long-run growth/welfare response**
  [static: Hsieh and Moretti (2019), Fajgelbaum and Gaubert (2020), Rossi-Hansberg, Sarte, and Schwartzman (2021)]

This paper: Three contributions

1. **Theory: Local human capital externalities $\rightarrow$ Agglomeration & Growth**
   - system of cities
   - heterogeneous workers **learn & migrate** over the life cycle
   - human capital process drives **both** agglomeration and growth
     - learn from others in your city, more if bigger or more skilled (*local externalities*)
     - learning $\rightarrow$ human capital dist. shifts right $\rightarrow$ output grows
   - **characterize "cities drive growth": growth rate = f(spatial distribution)**

This paper: Three contributions

1. **Theory: Local human capital externalities $\rightarrow$ Agglomeration & Growth**
   - system of cities
   - heterogeneous workers **learn & migrate** over the life cycle
   - human capital process drives **both** agglomeration and growth
     - learn from others in your city, more if bigger or more skilled (*local externalities*)
     - learning $\rightarrow$ human capital dist. shifts right $\rightarrow$ output grows
   - **characterize "cities drive growth": growth rate $=$ f(spatial distribution)**

Solves the **hard problem of regional econ** (Breinlich, Ottaviano, and Temple, 2014)

- "How to model growth and agglomeration as outcomes of a joint process"
- Agents must know *distribution* of economic activity over *time & space* $\rightarrow$ high-dimensional
- *how?* **Equilibrium is a mean field game** (Achdou et al., 2022) $\rightarrow$ can track distribution

This paper: Three contributions

2. **Quantitative: Using U.S. data, jointly rationalize . . .**
   - . . . **urban cross-section**:
     - match city size distribution
     - big cities more productive, more expensive, more skilled on avg. (Glaeser, 2008)
   - . . . **worker panels**:
     - life-cycle of human capital investment (Ben-Porath, 1967; Huggett, Ventura, and Yaron, 2006)
     - migration driven by expected income; young & edu. move more (Kennan and Walker, 2011)
     - city size wage premium = higher wage level + **faster wage growth w/ permanent value** (Glaeser and Maré, 2001; Baum-Snow and Pavan, 2012; Duranton and Puga, 2022)
   - . . . **aggregate growth**: 2% per year on BGP

This paper: Three contributions

3. **Long-run effects of place-based policy**
    - policy: relax LURs in **NY** and **SF** to U.S. median
    - outcome: aggregate growth **increases by 13bp**
    - **through what channel?**
        - *not* syphoning skill from elsewhere
        - instead, stronger dynamic spillover → **faster human capital accumulation**

This paper: Three contributions

3. **Long-run effects of place-based policy**
    - policy: relax LURs in **NY** and **SF** to U.S. median
    - outcome: aggregate growth **increases by 13bp**
    - **through what channel?**
        - *not* syphoning skill from elsewhere
        - instead, stronger dynamic spillover $\rightarrow$ **faster human capital accumulation**

Spatial **policy** $\rightarrow \Delta$ spatial **distribution** $\rightarrow \Delta$ **growth** in two (complementary) ways:

- by *attracting* more skilled workers to particular cities (e.g., push *skilled* to **NY**)
- by *producing* more skilled workers for the economy overall (e.g., push *young* to **NY**)

Outline for today

1. **Model**: setup, equilibrium, BGP, main result
2. **Quantitative analysis**: calibration/estimation, predictions
3. **Counterfactual place-based policy**
4. **Conclusion**

Model

Environment

- continuous time $t \in [0, \infty)$, discrete cities $n = 1, \ldots, N$
- mass L of workers with **human capital** $z \in \mathbb{R}_{++}$ and **age** $a \in [0, A]$
  - discount at rate ρ
  - hand-to-mouth
  - consume traded good c (numeraire) and land (**strict necessity**), benefit from amenity B_n

Environment

- continuous time $t \in [0, \infty)$, discrete cities $n = 1, \ldots, N$
- mass L of workers with **human capital** $z \in \mathbb{R}_{++}$ and **age** $a \in [0, A]$
  - discount at rate ρ
  - hand-to-mouth
  - consume traded good c (numeraire) and land (**strict necessity**), benefit from amenity B_n
- two **choices** at each t:
  - **raw labor**: learn (s) and work $(1 - s) \rightarrow$ income tomorrow vs. today
  - **migration**: city n s.t. opportunity $\overset{iid}{\sim}$ Poisson(λ) & taste $b_n^{\omega} \overset{iid}{\sim}$ T2EV(ϵ) & cost τ_{in}

Environment

- continuous time $t \in [0, \infty)$, discrete cities $n = 1, \ldots, N$
- mass L of workers with **human capital** $z \in \mathbb{R}_{++}$ and **age** $a \in [0, A]$
  - discount at rate ρ
  - hand-to-mouth
  - consume traded good c (numeraire) and land (**strict necessity**), benefit from amenity B_n
- two **choices** at each t:
  - **raw labor**: learn (s) and work $(1 - s)$ $\longrightarrow$ income tomorrow vs. today
  - **migration**: city n s.t. opportunity $\overset{iid}{\sim}$ Poisson(λ) & taste $b_n^\omega \overset{iid}{\sim}$ T2EV(ϵ) & cost τ_{in}
- expected flow utility:

$$U_n(s; z, t) = \underbrace{B_n}_{\textbf{Amenity}} \underbrace{[\overbrace{y_n(s; z, t)}^{\textbf{Income}} - \overbrace{P_n(z, t)}^{\textbf{Urban cost}}]}_{\textbf{Consumption flow}}$$

City characteristics: Congestion vs. Agglomeration

Endogenous city populations: $L_n(t) = L \iint g_n(a, z, t)\, dz\, da$

City characteristics: Congestion vs. Agglomeration

Endogenous city populations: $L_n(t) = L \iint g_n(a, z, t)\, dz\, da$

Congestion through land

- pay heterogeneous flow cost $\boldsymbol{z P_n(t)}$ for accomodation in city n, where

$$P_n(t) = p_n \boldsymbol{L_n(t)^{\theta_n}}$$

- *microfoundation*: monocentric city with commuting cost as forgone income (Duranton and Puga, 2015)

City characteristics: Congestion vs. Agglomeration

Endogenous city populations: $L_n(t) = L \iint g_n(a, z, t)\, dz\, da$

Congestion through land

- pay heterogeneous flow cost $\boldsymbol{zP_n(t)}$ for accomodation in city n, where

$$P_n(t) = p_n \boldsymbol{L_n(t)^{\theta_n}}$$

- *microfoundation*: monocentric city with commuting cost as forgone income (Duranton and Puga, 2015)

Agglomeration through static & dynamic channels

- Through **income**: $y_n(s; z, t) = T_n \boldsymbol{L_n(t)^{\alpha}} (1-s) z$

City characteristics: Congestion vs. Agglomeration

Endogenous city populations: $L_n(t) = L \iint g_n(a, z, t)\, dz\, da$

Congestion through land

- pay heterogeneous flow cost $\boldsymbol{zP_n(t)}$ for accomodation in city n, where

$$P_n(t) = p_n \boldsymbol{L_n(t)}^{\boldsymbol{\theta_n}}$$

- *microfoundation*: monocentric city with commuting cost as forgone income

(Duranton and Puga, 2015)

Agglomeration through static & dynamic channels

- Through **income**: $y_n(s; z, t) = T_n \boldsymbol{L_n(t)^{\alpha}}(1-s)z$
- Define the **vibrancy** of city n as

$$Z_n(t) = \left(L \iint z^{\zeta} g_n(a, z, t) dz\, da\right)^{\frac{1}{\zeta}}$$
$$\equiv \boldsymbol{L_n(t)}^{\boldsymbol{\frac{1}{\zeta}}} \bar{z}_{n,\zeta}(t)$$

- Through **learning**: law of motion for skill

$$\frac{dz}{dt} = \kappa(s) z^{\beta} Z_n(t)^{1-\beta}$$

Some notes on learning

$$\frac{dz}{dt} = \kappa(s) z^{\beta} Z_n(t)^{1-\beta}$$

where

- $Z_n(t) = \boldsymbol{L_n(t)}^{\frac{1}{\zeta}} \bar{z}_{n,\zeta}(t)$
- $\kappa(0) = 0, \ \kappa' > 0, \ \kappa'' < 0$

What to notice:

- **agglomeration: anyone can learn from anyone...**
- supermodularity
- classical form
- returns to scale

Some notes on learning

$$\frac{dz}{dt} = \kappa(s) \boldsymbol{z^{\beta} Z_n(t)^{1-\beta}}$$

where

- $Z_n(t) = L_n(t)^{\frac{1}{\zeta}} \boldsymbol{\bar{z}_{n,\zeta}(t)}$
- $\kappa(0) = 0,\ \kappa' > 0,\ \kappa'' < 0$

What to notice:

- agglomeration
- **supermodularity: . . . but more if/from highly-skilled**
- classical form
- returns to scale

Some notes on learning

$$\frac{dz}{dt} = \boldsymbol{\kappa(s) z^{\beta}} Z_n(t)^{1-\beta}$$

where

- $Z_n(t) = L_n(t)^{\frac{1}{\zeta}} \bar{z}_{n,\zeta}(t)$
- $\kappa(0) = 0,\ \kappa' > 0,\ \kappa'' < 0$

What to notice:

- agglomeration
- supermodularity
- **classical form**: **Ben-Porath (1967)**, **Rosen (1976)**, **Heckman (1976)**
- returns to scale

Some notes on learning

$$\frac{dz}{dt} = \boldsymbol{\kappa}(s) z^{\boldsymbol{\beta}} Z_n(t)^{\mathbf{1}-\boldsymbol{\beta}}$$

where

- $Z_n(t) = L_n(t)^{\frac{1}{\zeta}} \bar{z}_{n,\zeta}(t)$
- $\kappa(0) = 0,\ \kappa' > 0,\ \boldsymbol{\kappa'' < 0}$

What to notice:

- agglomeration
- supermodularity
- classical form
- **returns to scale: DRS in time, CRS in tuple (z, Z_n)**

Worker's problem

- Given city sizes $\mathbf{L}(t) \equiv \{L_1(t), \ldots, L_N(t)\}$ and vibrancies $\mathbf{Z}(t) \equiv \{Z_1(t), \ldots, Z_N(t)\}$
- **Hamilton-Jacobi-Bellman equation**: after expectations over T2EV preferences...

$$\rho V_n(a,z,t) = \max_{s\in[0,1]} \left\{ \underbrace{B_n[T_n L_n(t)^\alpha (1-s) - P_n(t)]z}_{\text{flow utility}} + \partial_z V_n(a,z,t) \underbrace{\left[\kappa(s) z^\beta Z_n(t)^{1-\beta}\right]}_{\text{skill gains}} \right\}$$

$$+ \underbrace{\lambda \sum_i m_{ni}(a,z,t) \left[\xi_{ni}(a,z,t) V_i(a,z,t) - V_n(a,z,t)\right]}_{\text{expected migration gains}}$$

$$+ \partial_a V_n(a,z,t) + \partial_t V_n(a,z,t)$$

with **optimal migration shares** and **selection effect**:

$$m_{ni}(a,z,t) = \frac{\tau_{ni}^{-\epsilon} V_i(a,z,t)^\epsilon}{\sum_k \tau_{nk}^{-\epsilon} V_k(a,z,t)^\epsilon} \qquad \xi_{ni}(a,z,t) = \frac{1}{N\tau_{ni}} m_{ni}(a,z,t)^{-\frac{1+\epsilon}{\epsilon}}$$

- **Terminal condition**: $V_n(A,z,t) = 0$ for all (n,z,t)

How do city characteristics evolve?

- **Recall:** $\{\mathbf{L}, \mathbf{Z}\}$ are entirely determined by the **distribution** of (a, z) across n
- **Demographics:** uniform marginal age distribution, entrants $\sim \underline{g}_n(z, t)$ replace exiters
- **Kolmogorov forward equation**:

$$\partial_t g_n(a,z,t) = -\underbrace{\partial_z[h_n(a,z,t)g_n(a,z,t)]}_{\text{skill accum.}} - \underbrace{\lambda[1-m_{nn}(a,z,t)]g_n(a,z,t)}_{\text{outflow: migration}}$$
$$+\underbrace{\lambda\sum_{i\neq n} m_{in}(a,z,t)g_i(a,z,t)}_{\text{inflow: migration}} - \underbrace{\partial_a g_n(a,z,t)}_{\text{aging}}$$

with initial condition $g_n(0,z,t) = \frac{1}{A}\underline{g}_n(z,t)$ and **optimal skill accumulation**:

$$h_n(a,z,t) = \kappa[s_n(a,z,t)]z^{\beta} Z_n(t)^{1-\beta}$$

Look for a balanced growth path

- An **equilibrium**: **HJB** (V, s, m) + **KFE** (g) + **feasibility** $(\{\boldsymbol{L}, \boldsymbol{Z}\})$

Look for a balanced growth path

- An **equilibrium**: **HJB** $(V,\ s,\ m)$ + **KFE** (g) + **feasibility** $(\{\boldsymbol{L}, \boldsymbol{Z}\})$
- A **balanced growth path** is a number γ and functions (v, σ, μ, ϕ) on $\mathcal{N} \times \mathcal{X} \times [0, A]$ s.t.

$$\begin{aligned} V_n(a,z,t) &= e^{\gamma t} \quad v_n\,(a,x) \\ s_n(a,z,t) &= \qquad \sigma_n\,(a,x) \\ m_{ni}(a,z,t) &= \qquad \mu_{ni}\,(a,x) \\ g_n(a,z,t) &= e^{-\gamma t}\, \phi_n\,(a,x) \end{aligned}$$

and (V, s, m, g) is an equilibrium with initial condition $g_n(a, z, 0) = \phi_n(a, z)$, where

$x \equiv ze^{-\gamma t}$ is **relative human capital**.

Look for a balanced growth path

- An **equilibrium**: **HJB** (V, s, m) + **KFE** (g) + **feasibility** ($\{\boldsymbol{L}, \boldsymbol{Z}\}$)

- A **balanced growth path** is a number γ and functions (v, σ, μ, ϕ) on $\mathcal{N} \times \mathcal{X} \times [0, A]$ s.t.

$$\begin{aligned} V_n(a,z,t) &= e^{\gamma t} \quad v_n\,(a,x) \\ s_n(a,z,t) &= \qquad \sigma_n\,(a,x) \\ m_{ni}(a,z,t) &= \qquad \mu_{ni}\,(a,x) \\ g_n(a,z,t) &= e^{-\gamma t}\,\phi_n\,(a,x) \end{aligned}$$

  and (V, s, m, g) is an equilibrium with initial condition $g_n(a, z, 0) = \phi_n(a, z)$, where

  $x \equiv ze^{-\gamma t}$ is **relative human capital**.

- **Needs:** l.b. and quantiles of $\underline{g}$ grow with those of g $\rightarrow$ **entrants getting better**

Look for a balanced growth path

- An **equilibrium**: **HJB** $(V,\ s,\ m)$ + **KFE** (g) + **feasibility** $(\{\boldsymbol{L}, \boldsymbol{Z}\})$
- A **balanced growth path** is a number γ and functions (v, σ, μ, ϕ) on $\mathcal{N} \times \mathcal{X} \times [0, A]$ s.t.

$$
\begin{aligned}
V_n(a,z,t) &= e^{\gamma t} \quad v_n\,(a,x) \\
s_n(a,z,t) &= \qquad \sigma_n\,(a,x) \\
m_{ni}(a,z,t) &= \qquad \mu_{ni}\,(a,x) \\
g_n(a,z,t) &= e^{-\gamma t}\,\phi_n\,(a,x)
\end{aligned}
$$

  and (V, s, m, g) is an equilibrium with initial condition $g_n(a,z,0) = \phi_n(a,z)$, where

  $x \equiv ze^{-\gamma t}$ is **relative human capital**.

- **Needs:** l.b. and quantiles of $\underline{g}$ grow with those of $g \rightarrow$ **entrants getting better**
- **Implies:** **constant** city size, **same** productivity growth (*in progress*: solving transitions)

The detrended BGP: number γ and functions (v, σ, μ, ϕ)

- workers solve a **detrended HJB equation**

$$
\begin{aligned}
(\rho - \gamma) v_n(a,x) = {} & B_n(T_n L_n^{\alpha}[1 - \sigma_n(a,x)] - p_n L_n^{\theta_n}) x + \partial_x v_n(a,x)[h_n(a,x) - \gamma \boldsymbol{x}] \\
& + \partial_a v_n(a,x) + \lambda \sum_i \mu_{ni}(a,x)[\xi_{ni}(a,x) v_i(a,x) - v_n(a,x)]
\end{aligned}
$$

The detrended BGP: number γ and functions $(\boldsymbol{v}, \boldsymbol{\sigma}, \boldsymbol{\mu}, \boldsymbol{\phi})$

- workers solve a **detrended HJB equation**

$$(\rho - \gamma) v_n(a,x) = B_n(T_n L_n^{\alpha}[1-\sigma_n(a,x)] - p_n L_n^{\theta_n})x + \partial_x v_n(a,x)[h_n(a,x) - \gamma \boldsymbol{x}] + \partial_a v_n(a,x) + \lambda \sum_i \mu_{ni}(a,x)[\xi_{ni}(a,x) v_i(a,x) - v_n(a,x)]$$

- the detrended distribution evolves according to a **Kolmogorov forward equation**

$$0 = -\partial_x\{[h_n(a,x) - \gamma \boldsymbol{x}]\phi_n(a,x)\} - \partial_a \phi_n(a,x) - \lambda[1-\mu_{nn}(a,x)]\phi_n(a,x) + \lambda \sum_{i \neq n} \mu_{in}(a,x)\phi_i(a,x)$$

The detrended BGP: number γ and functions $(\boldsymbol{v}, \boldsymbol{\sigma}, \boldsymbol{\mu}, \boldsymbol{\phi})$

- workers solve a **detrended HJB equation**

$$
\begin{aligned}
(\rho - \gamma) v_n(a,x) = {} & B_n(T_n L_n^{\alpha}[1-\sigma_n(a,x)] - p_n L_n^{\theta_n})x + \partial_x v_n(a,x)[h_n(a,x) - \gamma \boldsymbol{x}] \\
& + \partial_a v_n(a,x) + \lambda \sum_i \mu_{ni}(a,x)[\xi_{ni}(a,x) v_i(a,x) - v_n(a,x)]
\end{aligned}
$$

- the detrended distribution evolves according to a **Kolmogorov forward equation**

$$
\begin{aligned}
0 = {} & -\partial_x\{[h_n(a,x) - \gamma \boldsymbol{x}]\phi_n(a,x)\} - \partial_a \phi_n(a,x) \\
& - \lambda[1-\mu_{nn}(a,x)]\phi_n(a,x) + \lambda \sum_{i \neq n} \mu_{in}(a,x)\phi_i(a,x)
\end{aligned}
$$

- aggregates are **feasible**:

$$
X_n = \left(L \iint x^{\zeta}\, \phi_n(a,x)\, dx\, da \right)^{\frac{1}{\zeta}} \qquad L_n = L \iint \phi_n(a,x)\, dx\, da
$$

The detrended BGP: number γ and functions (v, σ, μ, ϕ)

- workers solve a **detrended HJB equation**

$$
\begin{aligned}
(\rho - \gamma) v_n(a,x) = {} & B_n(T_n L_n^{\alpha}[1-\sigma_n(a,x)] - p_n L_n^{\theta_n})x + \partial_x v_n(a,x)[h_n(a,x) - \gamma x] \\
& + \partial_a v_n(a,x) + \lambda \sum_i \mu_{ni}(a,x)[\xi_{ni}(a,x) v_i(a,x) - v_n(a,x)]
\end{aligned}
$$

- the detrended distribution evolves according to a **Kolmogorov forward equation**

$$
\begin{aligned}
0 = {} & -\partial_x\{[h_n(a,x) - \gamma x]\phi_n(a,x)\} - \partial_a \phi_n(a,x) \\
& - \lambda[1-\mu_{nn}(a,x)]\phi_n(a,x) + \lambda \sum_{i \neq n} \mu_{in}(a,x)\phi_i(a,x)
\end{aligned}
$$

- aggregates are **feasible**:

$$
X_n = \left(L \iint x^{\zeta}\, \phi_n(a,x)\, dx\, da \right)^{\frac{1}{\zeta}} \qquad L_n = L \iint \phi_n(a,x)\, dx\, da
$$

- an expression relating the **growth rate** to the **decision rules & distribution**

Main theoretical result: Endogenous growth rate $\gamma(\sigma, \phi)$

Add up KF over all (n, a, x), noting **no net migration** and **uniform age density**:

$$\gamma(\sigma, \phi) = \frac{\sum_n \int \kappa[\sigma_n(a, x)] x^\beta X_n^{1-\beta} \phi_n(a, x)\, da}{\sum_n \int x \phi_n(a, x)\, da}, \quad \forall x \in \text{supp}(\phi)$$

Main theoretical result: Endogenous growth rate $\gamma(\sigma, \phi)$

Add up KF over all (n, a, x), noting **no net migration** and **uniform age density**:

$$\gamma(\sigma, \phi) = \frac{\sum_n \int \kappa[\sigma_n(a,x)] x^\beta X_n^{1-\beta} \phi_n(a,x)\, da}{\sum_n \int x \phi_n(a,x)\, da}, \quad \forall x \in \mathrm{supp}(\phi)$$

Intuition by building on Uzawa (1965):

$$\max_s \int_0^\infty e^{-\rho t} c\, dt \quad \text{s.t.} \quad \begin{array}{c} c = (1-s)z \\ \dot{z} = \kappa(s) z \end{array} \quad \Longrightarrow \quad \gamma = \kappa(s^*)$$

Main theoretical result: Endogenous growth rate $\gamma(\sigma, \phi)$

Add up KF over all (n, a, x), noting **no net migration** and **uniform age density**:

$$\gamma(\sigma, \phi) = \frac{\sum_n \int \kappa[\sigma_n(a, x)] x^\beta X_n^{1-\beta} \phi_n(a, x)\, da}{\sum_n \int x \phi_n(a, x)\, da}, \quad \forall x \in \mathsf{supp}(\phi)$$

Intuition by building on Uzawa (1965): reintroduce externality ...

$$\max_s \int_0^\infty e^{-\rho t} c\, dt \quad \text{s.t.} \quad \begin{aligned} c &= (1-s)z \\ \dot{z} &= \kappa(s) z^\beta Z^{1-\beta} \end{aligned} \quad \Longrightarrow \gamma = \kappa(s_1^*) < \kappa(s^*)$$

Main theoretical result: Endogenous growth rate $\gamma(\sigma, \phi)$

Add up KF over all (n, a, x), noting **no net migration** and **uniform age density**:

$$\gamma(\sigma, \phi) = \frac{\sum_n \int \kappa[\sigma_n(a, x)] x^\beta X_n^{1-\beta} \phi_n(a, x)\, da}{\sum_n \int x \phi_n(a, x)\, da}, \quad \forall x \in \mathsf{supp}(\phi)$$

Intuition by building on Uzawa (1965): ...then reintroduce death (random at rate δ)

$$\max_s \int_0^\infty e^{-(\rho+\delta)t} c\, dt \quad \text{s.t.} \quad \begin{array}{c} c = (1-s)z \\ \dot{z} = \kappa(s) z^\beta Z^{1-\beta} \end{array} \implies \gamma = \kappa(s_2^*) < \kappa(s_1^*)$$

Main theoretical result: Endogenous growth rate $\gamma(\sigma, \phi)$

Add up KF over all (n, a, x), noting **no net migration** and **uniform age density**:

$$\gamma(\sigma, \phi) = \frac{\sum_n \int \kappa[\sigma_n(a,x)] x^\beta X_n^{1-\beta} \phi_n(a,x)\, da}{\sum_n \int x \phi_n(a,x)\, da}, \quad \forall x \in \mathsf{supp}(\phi)$$

Intuition by building on Uzawa (1965): . . . then reintroduce worker heterogeneity (a, x)

$$\gamma = \frac{\int_0^A \boldsymbol{\kappa[\sigma(a,x)]} x^\beta X^{1-\beta}\, \phi(a,x)\, da}{\int_0^A x\, \phi(a,x)\, da}, \quad \forall x \in \mathsf{supp}(\phi) \ \text{ with } \ X = \left(L \iint x^\zeta \phi(a,x)\, dx\, da \right)^{\frac{1}{\zeta}}$$

Main theoretical result: Endogenous growth rate $\gamma(\sigma, \phi)$

Add up KF over all (n, a, x), noting **no net migration** and **uniform age density**:

$$\gamma(\sigma, \phi) = \frac{\sum_n \int \kappa[\sigma_n(a,x)] x^\beta X_n^{1-\beta} \phi_n(a,x)\, da}{\sum_n \int x \phi_n(a,x)\, da}, \quad \forall x \in \text{supp}(\phi)$$

Intuition by building on Uzawa (1965): . . . then, finally, reintroduce space

- growth rate = weighted avg. of returns to investment, $\kappa(\sigma)$
- more weight to investment in **larger, more skilled places**
- **spatial distribution of human capital matters for growth**

Main theoretical result: Endogenous growth rate $\gamma(\sigma, \phi)$

Add up KF over all (n, a, x), noting **no net migration** and **uniform age density**:

$$\gamma(\sigma, \phi) = \frac{\sum_n \int \kappa[\sigma_n(a,x)] x^\beta X_n^{1-\beta} \phi_n(a,x)\, da}{\sum_n \int x \phi_n(a,x)\, da}, \quad \forall x \in \text{supp}(\phi)$$

Intuition by building on Uzawa (1965): ...then, finally, reintroduce space

- growth rate = weighted avg. of returns to investment, $\kappa(\sigma)$
- more weight to investment in **larger, more skilled places**
- **spatial distribution of human capital matters for growth**

When every idea must be in somebody's brain, it matters where those brains are.

Recap: How it solves the "hard problem"

Key idea: Economy is summarized by density $g_n(a, z, t)$, which we can track!

Recap: How it solves the “hard problem”

Key idea: Economy is summarized by density $g_n(a, z, t)$, which we can track!

Look for a **BGP**, which is characterized by three equations (**MFG** + **growth**): ▸ well-posedness

1. **HJB**: how workers w/ states (n, a, z) learn and migrate given (moments of) distribution
2. **Kolmogorov forward**: how distribution evolves in response to workers' decisions
3. **endogenous growth**: restriction on distribution relating cross-sectional **shape** to **speed**

$$\textbf{growth rate} = f(\textbf{spatial distribution})$$

Recap: How it solves the "hard problem"

Key idea: Economy is summarized by density $\boldsymbol{g_n(a, z, t)}$, which we can track!

Look for a **BGP**, which is characterized by three equations (**MFG** + **growth**): ▸ well-posedness

1. **HJB**: how workers w/ states (n, a, z) learn and migrate given (moments of) distribution
2. **Kolmogorov forward**: how distribution evolves in response to workers' decisions
3. **endogenous growth**: restriction on distribution relating cross-sectional **shape** to **speed**

$$\textbf{growth rate} = f(\textbf{spatial distribution})$$

Spatial **policy** $\rightarrow \Delta$ spatial **distribution** $\rightarrow \Delta$ **growth** in two (complementary) ways:

- by *attracting* more skilled workers to particular cities (e.g., send higher $\boldsymbol{x}$ to raise $\boldsymbol{X_n}$)
- by *producing* more skilled workers throughout (e.g., send higher $\boldsymbol{\sigma}_n$ to already-high $\boldsymbol{X_n}$)

Quantitative analysis

Four steps to rationalize patterns in U.S. data

Select cities:

- 378 MSAs $\rightarrow$ **30 biggest + 4 groups**
- congestion elasticity from (Saiz, 2010)

Four steps to rationalize patterns in U.S. data

Select cities:

- 378 MSAs $\rightarrow$ **30 biggest + 4 groups**
- congestion elasticity from (Saiz, 2010)

Quantify parameters in three groups:

1. Migration frictions: Back out from 2011–15 ACS migration flows cross-tabbed by age

   $\star$ target mobility by age & avg. bilateral likelihood, test against bilateral patterns by age

Four steps to rationalize patterns in U.S. data

Select cities:

- 378 MSAs $\rightarrow$ **30 biggest + 4 groups**
- congestion elasticity from (Saiz, 2010)

Quantify parameters in three groups:

1. Migration frictions: Back out from 2011–15 ACS migration flows cross-tabbed by age
   - $\star$ target mobility by age & avg. bilateral likelihood, test against bilateral patterns by age
2. Life-cycle investment: Calibrate to structural estimates of Ben-Porath model (nested)

Four steps to rationalize patterns in U.S. data

Select cities:

- 378 MSAs $\rightarrow$ **30 biggest + 4 groups**
- congestion elasticity from (Saiz, 2010)

Quantify parameters in three groups:

1. Migration frictions: Back out from 2011–15 ACS migration flows cross-tabbed by age

   $\star$ target mobility by age & avg. bilateral likelihood, test against bilateral patterns by age

2. Life-cycle investment: Calibrate to structural estimates of Ben-Porath model (nested)

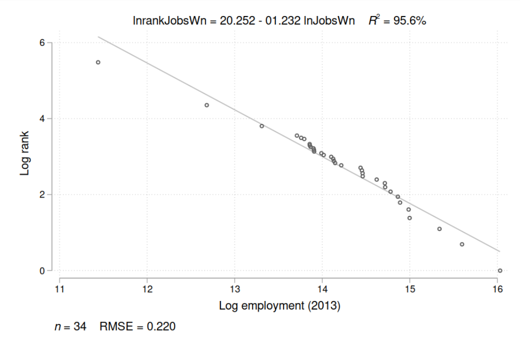
   $\star$ (computational and data constraints prevent folding #2 into #3 for now)

3. Local fundamentals and elasticities: MDE to match data, prior lit.
   - B_n, T_n, p_n: city size, income; budget shares
   - α, ζ, x_{scale}: Duranton and Puga (2022) wage regression on NLSY panel w/ city groups

Using U.S. data, the model can jointly rationalize...

...urban cross-section:

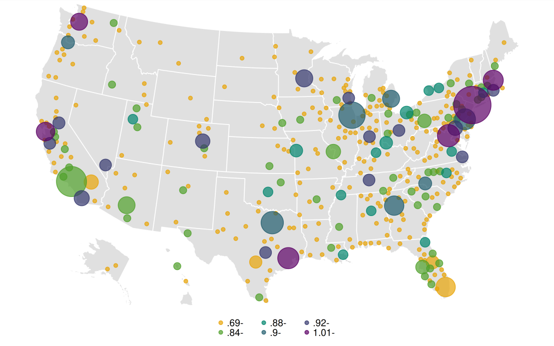
- match city size distribution
- big cities more productive, more expensive, more skilled on avg. (Glaeser, 2008)



Using U.S. data, the model can jointly rationalize. . .

. . . **urban cross-section:**

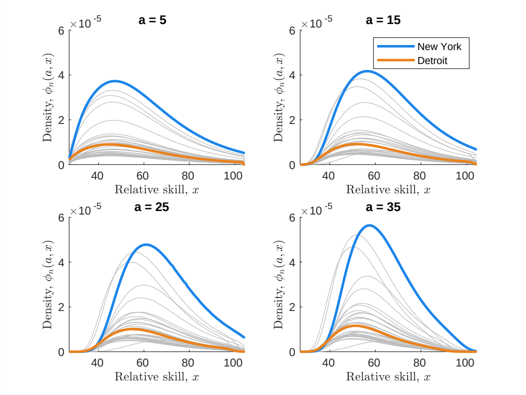
- match city size distribution
- big cities more productive, more expensive, more skilled on avg. (Glaeser, 2008)



Using U.S. data, the model can jointly rationalize...

...urban cross-section:

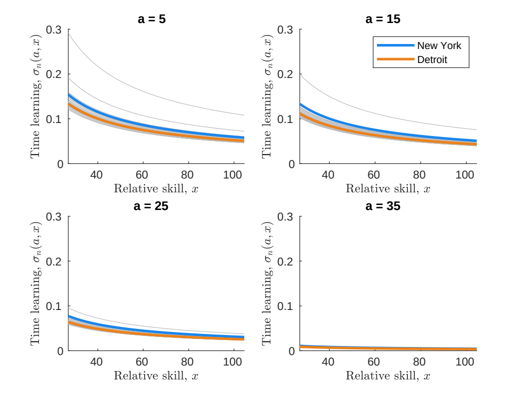
- match city size distribution
- big cities more productive, more expensive, more skilled on avg. (Glaeser, 2008)



Using U.S. data, the model can jointly rationalize. . .

. . . worker panels:

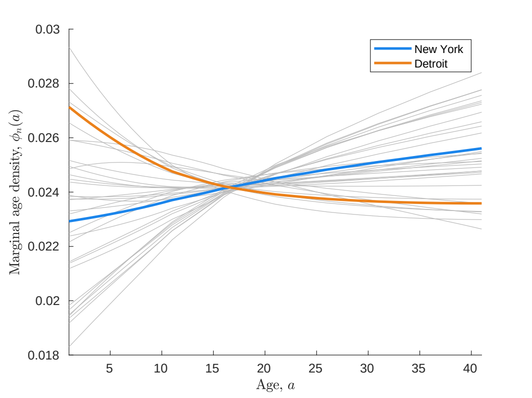
- life-cycle of human capital investment (Ben-Porath, 1967; Huggett, Ventura, and Yaron, 2006)
- migration driven by expected income; young & educated move more (Kennan and Walker, 2011)
- city size wage premium = higher wage level + **faster wage growth w/ permanent value** (Glaeser and Maré, 2001; Baum-Snow and Pavan, 2012; Duranton and Puga, 2022)



Using U.S. data, the model can jointly rationalize. . .

. . . worker panels:

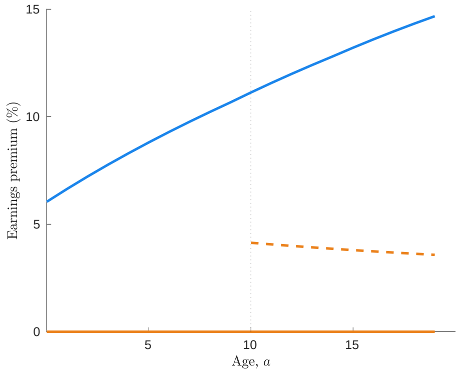
- life-cycle of human capital investment (Ben-Porath, 1967; Huggett, Ventura, and Yaron, 2006)
- migration driven by expected income; young & educated move more (Kennan and Walker, 2011)
- city size wage premium = higher wage level + **faster wage growth w/ permanent value** (Glaeser and Maré, 2001; Baum-Snow and Pavan, 2012; Duranton and Puga, 2022)



Using U.S. data, the model can jointly rationalize. . .

. . . worker panels:

- life-cycle of human capital investment (Ben-Porath, 1967; Huggett, Ventura, and Yaron, 2006)
- migration driven by expected income; young & educated move more (Kennan and Walker, 2011)
- city size wage premium = higher wage level + **faster wage growth w/ permanent value** (Glaeser and Maré, 2001; Baum-Snow and Pavan, 2012; Duranton and Puga, 2022)



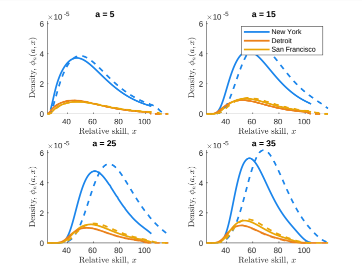
Using U.S. data, the model can jointly rationalize...

...**aggregate growth:** 2% per year on BGP

Policy counterfactual

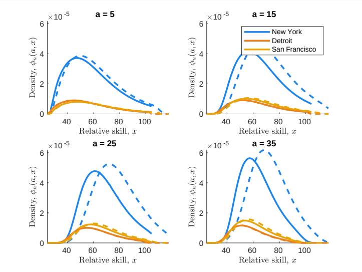
Counterfactual: Relaxing LURs in NY and SF

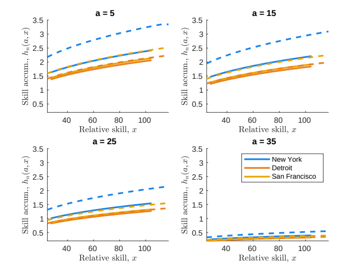
- Relax land use regulations in **NY** and **SF** to median level ($\downarrow \theta_n$)
- On new BGP, both cities would have lower urban costs at old pop. levels
  - **direct:** lower costs attract workers
  - **indirect:** static & dynamic agglomeration amplify attraction
  - **result:** both *bigger*, *more skilled*



Counterfactual: Relaxing LURs in NY and SF

- Relax land use regulations in **NY** and **SF** to median level ($\downarrow \theta_n$)
- On new BGP, both cities would have lower urban costs at old pop. levels
  - **direct:** lower costs attract workers
  - **indirect:** static & dynamic agglomeration amplify attraction
  - **result:** both *bigger*, *more skilled*
- **Through what channel?**
  - *not* syphoning skill from elsewhere
  - instead, stronger dynamic spillover $\rightarrow$ **faster human capital accumulation**
- Overall, **growth** $\uparrow$**13bp** b/c policy **produces** more skilled workers





Recap: The "hard problem" of regional economics

- **A classic hypothesis** (Jacobs, 1969; Lucas, 1988; Glaeser, 2011) **. . .**

	human capital spillovers	$\rightarrow$ agglomeration (cities!)
$+$	human capital accumulation	$\rightarrow$ growth
	human capital accumulation s.t. local spillovers	$\rightarrow$ **"cities drive growth"**

. . . but no models $\rightarrow$ no testing, no counterfactuals, no optimal policy

- **Why not?** forward-looking dynamics + can't average across space $\rightarrow$ *high-dimensional*

- **This paper: Tackle hard problem with new tools + apply to U.S. data & policy**
  1. characterize "cities drive growth": growth rate $= f$(spatial distribution)
  2. rationalize patterns in U.S. data: worker panels, city cross-section, aggregate BGP trend
  3. policy counterfactual: relax LURs in **NY** and **SF** $\rightarrow$ aggregate growth increases by **13bp**

Appendix

Formal definition of location history

Key: taste shocks and migrations costs for chosen cities are *multiplicative* and *permanent* (Desmet, Nagy, and Rossi-Hansberg, 2018; Caliendo, Dvorkin, and Parro, 2019)

- birth location n_0^ω
- count of opportunities $I^\omega(t)$, a Poisson process with arrival rate λ
- each opportunity ι, draw idiosyncratic taste shocks $\{b_n^{\omega,\iota}\}_n \overset{iid}{\sim}$ T2EV(ϵ)
- bilateral migration costs τ_{ni}
- **define . . .**
  - $n_\iota^\omega :=$ her location choice at opportunity ι
  - $\hat{b}_\iota^\omega := b_{n_\iota^\omega}^{\omega,\iota}$, the realization of her taste shock for her location choice
  - $\hat{\tau}_\iota^\omega := \tau_{n_{\iota-1}^\omega, n_\iota^\omega}$, the bilateral cost to move to her location choice
- **then we have . . .**

$$U^\omega(t) = \Omega^\omega(t) U_n(s; z, t) \text{ with } \Omega^\omega(t) := \prod_{\iota=1}^{I^\omega(t)} \frac{\hat{b}_\iota^\omega}{\hat{\tau}_\iota^\omega}$$

City structure and urban costs: Setup

- canonical rent gradient model: trade off commuting cost vs. rents, utility equalizes
- **here**: heterogeneous agents $\implies$ assignment problem
- a city is a line with...
  - all production at single point ("CBD")
  - identical residences of unit length
- **commuting takes time**: forgo $(T_n z)\vartheta_n \ell^\theta$ of income to commute from distance ℓ
- **equilibrium**: a rent gradient $r_n(\ell, t)$ and an assignment function $\mathcal{L}_n(z, t)$ s.t.
  (i) individual optimality holds (Alonso-Muth condition):

$$\theta(T_n z)\vartheta_n \mathcal{L}_n(z,t)^\theta = -\partial_\ell r_n(\mathcal{L}_n(z,t), t)$$

  (ii) all workers are allocated to a residence

City structure and urban costs: Solving for equilibrium

- "supply = demand" + sorting: **allocate by skill quantile**

$$-\frac{\partial \mathcal{L}_n(z,t)}{\partial z} = \frac{1}{2} g_n(z,t) \implies \mathcal{L}_n(z,t) = \frac{L_n(t)}{2}[1 - G_n(z,t)]$$

  where $g_n(z,t)$ is the marginal density of skill (integrated over age)

- **rents**: integrate Alonso-Muth condition given assignment function

$$r_n(\ell,t) = \theta \vartheta_n T_n \int_{\ell}^{L_n(t)/2} G_n^{-1}\left(1 - \frac{2l}{L_n(t)}, t\right) l^{\theta-1} dl.$$

- urban cost grows at the same rate as income $\implies$ **consumption grows at constant rate**
- would need to guess G each iteration $\implies$ let local government collect & redistribute rents to simplify urban cost to $zP_n(t) = \theta \vartheta_n T_n z L_n(t)^{\theta}$

Equilibrium = Mean Field Game (MFG)

A tuple of functions $\{V, s, m, g\}$ on $\mathcal{N} \times \mathbb{R}_{++} \times [0, A] \times \mathbb{R}_+$ and a tuple of functions $\{\mathbf{L}, \mathbf{P}, \mathbf{Z}\}$ on $\mathcal{N} \times \mathbb{R}_+$ such that

1. workers solve the **Hamilton-Jacobi-Bellman equation** for $n = 1, \ldots, N$, taking paths of vibrancies $\mathbf{Z}$ and city sizes $\mathbf{L}$ (thus, also urban costs $\mathbf{P}$) as given;
2. density $g_n(a, z, t)$ evolves according to the **Kolmogorov forward equation** for $n = 1, \ldots, N$, taking workers' optimal policy functions as given;
3. vibrancies and urban costs **satisfy their definitions** given $g_n(a, z, t)$:

$$Z_n(t) = \left(L \iint z^\zeta g_n(a, z, t)\, dz\, da \right)^{\frac{1}{\zeta}}, \qquad P_n(t) = p_n \left(L \iint g_n(a, z, t)\, dz\, da \right)^{\theta_n};$$

4. local population shares **sum to one** for all t:

$$1 = \sum_{n=1}^{N} \frac{L_n(t)}{L} = \sum_{n=1}^{N} \iint g_n(a, z, t)\, dz\, da.$$

Loose end: Well-posedness of a BGP?

- Main result: "if BGP $\{v, \sigma, \mu, \phi, \gamma\}$ exists, it must be that $\gamma = f(\sigma, \phi)$"

Loose end: Well-posedness of a BGP?

- Main result: "if BGP $\{v, \sigma, \mu, \phi, \gamma\}$ exists, it must be that $\gamma = f(\sigma, \phi)$"
- **But does a BGP exist?** **no proofs** of existence & uniqueness for this class of MFGs
  (first-order, smoothing, non-separable Hamiltonian w/ boundaries)

Loose end: Well-posedness of a BGP?

- Main result: "if BGP $\{v, \sigma, \mu, \phi, \gamma\}$ exists, it must be that $\gamma = f(\sigma, \phi)$"
- **But does a BGP exist?** **no proofs** of existence & uniqueness for this class of MFGs (first-order, smoothing, non-separable Hamiltonian w/ boundaries)
- take **three steps** to ensure sensible results:
  1. impose $\beta\bar{\kappa} < \rho\eta$, necessary for E&U in Uzawa model with externality
  2. show **existence & local stability of (discretized) BGP by construction** $\rightarrow$ look for one that matches data
  3. in counterfactual, select closest BGP (Ahlfeldt et al., 2015)

Overview of algorithm

Adapt the usual **HACT algorithm** (Achdou et al., 2022)

0. Begin with guess $\{\gamma^0, \mathbf{X}^0, \mathbf{L}^0\}$. Denote iterations by $\iota = 0, 1, 2, \ldots$.
1. Given $\{\gamma^\iota, \mathbf{X}^\iota, \mathbf{L}^\iota\}$, solve detrended HJB w/ finite difference method + calculate policy functions $\sigma_n^\iota(a,x)$ and $\mu_{ni}^\iota(a,x)$.
2. Given $\sigma_n^\iota(a,x)$ and $\mu_{ni}^\iota(a,x)$, solve KF for $\phi_n^\iota(a,x)$ w/ finite difference method.
3. Given $\phi_n^\iota(a,x)$, compute vibrancies, populations, and growth rate:

$$\tilde{X}_n^\iota = \left(L \iint x^\zeta \, \phi_n^\iota(a,x) \, dx \, da \right)^{\frac{1}{\zeta}}, \qquad \tilde{L}_n^\iota = L \iint \phi_n^\iota(a,x) \, dx \, da.$$

$$\tilde{\gamma}^\iota = \frac{\sum_n \int_0^A \kappa[\sigma_n^\iota(a,x)] x^\beta (\tilde{X}_n^\iota)^{1-\beta} \, \phi_n^\iota(a,x) \, da}{\sum_n \int_0^A x \, \phi_n^\iota(a,x) \, da}, \quad x \in \text{supp}(\phi^\iota)$$

4. If $\{\tilde{\gamma}^\iota, \tilde{\mathbf{X}}^\iota, \tilde{\mathbf{L}}^\iota\}$ close enough to $\{\gamma^\iota, \mathbf{X}^\iota, \mathbf{L}^\iota\}$, **stop**. Else, construct $\{\gamma^{\iota+1}, \mathbf{X}^{\iota+1}, \mathbf{L}^{\iota+1}\}$ as a linear combination of previous guess and computed values, then return to step 1.

One-slide summary of steps 1 & 2

- Will discretize and solve using a **finite difference method** to approx. derivatives
- Discretization $\rightarrow$ HJB non-linear in $\mathbf{v}$, KF linear in ϕ, solved iteratively over age index j

$$(\rho-\gamma)\mathbf{v}^j = \mathbf{u}(\mathbf{v}^{j+1}) + \mathbf{\Pi}(\mathbf{v}^{j+1})\mathbf{v}^j \tag{HJB}$$

$$\mathbf{0} = \left(\mathbf{\Pi}^j\right)^T \boldsymbol{\phi}^j - \frac{\boldsymbol{\phi}^{j+1}-\boldsymbol{\phi}^j}{\Delta a} \tag{KF}$$

where each $\mathbf{\Pi}^j$ is a **sparse** transition matrix (rows sum to one)

Finite difference approximations to $v_n'(x_i)$

- Approximate $v_n(a,x)$ at $I \times J$ discrete points in the state space, x_i, $i = 1, \dots, I$, and a_j, $j = 1, \dots, J$ with distance Δx and Δa between points, resp.
- Shorthand notation: $v_{i,n}^j := v_n(a_j, x_i)$
- Need to approximate $\partial_x v_n(a_j, x_i)$ and $\partial_a v_n(a_j, x_i)$
- **Three different possibilities**: written for x, analogous for a

$$\partial_x^F v_{i,n}^j := \frac{v_{i+1,n}^j - v_{i,n}^j}{\Delta x} \qquad \text{forward difference}$$

$$\partial_x^B v_{i,n}^j := \frac{v_{i,n}^j - v_{i-1,n}^j}{\Delta x} \qquad \text{backward difference}$$

$$\partial_x^C v_{i,n}^j := \frac{v_{i+1,n}^j - v_{i-1,n}^j}{2\Delta x} \qquad \text{central difference}$$

Which to use? Always upwind!

- Best solution: **upwind scheme**
  - **forward** difference whenever drift of state variable is **positive**
  - **backward** difference whenever drift of state variable is **negative**
- Upwind version of HJB:

$$(\rho-\gamma)v^j_{i,n} = u^j_{i,n} + \partial^F_x v^j_{i,n}[h^j_{i,n} - \gamma x_i]^+ + \partial^B_x v^j_{i,n}[h^j_{i,n} - \gamma x_i]^- + \partial^F_a v^j_{i,n} + \lambda \sum_k \mu^j_{i,nk}[v^j_{i,k} - v^j_{i,n}]$$

with $y^+ = \max\{y, 0\}$ and $y^- = \min\{y, 0\}$ for any y

- **Complication**: drift $d^j_{i,n} \equiv h^j_{i,n} - \gamma x_i$ itself depends on which approx. is used

$$h^j_{i,n} = \kappa(\sigma^j_{i,n}) x_i^\beta X_n^{1-\beta}, \quad \text{where } \sigma^j_{i,n} \text{ is a function of } \partial_x v^j_{i,n}$$

- **Solution**: use $\sigma^{F,j}_{i,n}$ and $h^{F,j}_{i,n}$ when drift is positive; use $\sigma^{B,j}_{i,n}$ and $h^{B,j}_{i,n}$ when negative

Constructing the transition matrix Π^j

- Stack the discretized age-a_j value functions into a column vector of length NI

$$\mathbf{v}^j = [v^j_{1,1}, \dots, v^j_{I,1}, v^j_{1,2}, \dots, v^j_{I,2}, \dots, v^j_{1,N}, \dots, v^j_{I,N},]'$$

- Define the matrix entries

$$\pi^{B,j}_{i,n} = -\frac{(h^{B,j}_{i,n} - \gamma x_i)^-}{\Delta x}$$
$$\pi^{F,j}_{i,n} = \frac{(h^{F,j}_{i,n} - \gamma x_i)^+}{\Delta x}$$
$$\tilde{\pi}^j_{i,n} = -\pi^{F,j}_{i,n} + \pi^{B,j}_{i,n} - \lambda[1 - \mu^j_{i,n}\xi^j_{i,n}]$$

- Will be $NI \times NI$, block tri-diagonal, rows sum to one, very sparse

$$
\tilde{\mathbf{\Pi}}^j = \begin{bmatrix}
\tilde{\boldsymbol{\pi}}_1^j & \tilde{\mathbf{M}}_2^j & \tilde{\mathbf{M}}_3^j & \cdots & \tilde{\mathbf{M}}_N^j \\
\tilde{\mathbf{M}}_1^j & \tilde{\boldsymbol{\pi}}_2^j & \tilde{\mathbf{M}}_3^j & \cdots & \tilde{\mathbf{M}}_N^j \\
\vdots & \tilde{\mathbf{M}}_2^j & \ddots & \ddots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \tilde{\mathbf{M}}_N^j \\
\tilde{\mathbf{M}}_1^j & \tilde{\mathbf{M}}_2^j & \cdots & \tilde{\mathbf{M}}_{N-1}^j & \tilde{\boldsymbol{\pi}}_N^j
\end{bmatrix}
$$

$$
\mathbf{M}_n^j = \begin{bmatrix}
\lambda\mu_{1,n}^j\xi_{1,n}^j & 0 & \cdots & \cdots & \cdots & 0 \\
0 & \lambda\mu_{2,n}^j\xi_{2,n}^j & 0 & \ddots & \ddots & \vdots \\
\vdots & 0 & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & \cdots & \cdots & 0 & \lambda\mu_{I,n}^j\xi_{I,n}^j
\end{bmatrix}
$$

$$
\boldsymbol{\pi}_n^j = \begin{bmatrix}
\tilde{\pi}_{1,1}^j & \pi_{1,1}^{F,j} & 0 & \cdots & \cdots & 0 \\
\pi_{2,1}^{B,j} & \tilde{\pi}_{2,1}^j & \pi_{2,1}^{F,j} & 0 & \ddots & \vdots \\
0 & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & 0 & \pi_{I-1,1}^{B,j} & \tilde{\pi}_{I-1,1}^j & \pi_{I-1,1}^{F,j} \\
0 & \cdots & \cdots & 0 & \pi_{I,1}^{B,j} & \tilde{\pi}_{I,1}^j
\end{bmatrix}
$$

Implicit method for HJB

- Solve HJB **iteratively** backwards from terminal condition $\mathbf{v}^J = \mathbf{0}$.
- Want to solve

$$(\rho - \gamma)\boldsymbol{v}^j = \mathbf{u}^{j+1} + \tilde{\mathbf{\Pi}}^{j+1}\mathbf{v}^j + \frac{\mathbf{v}^{j+1} - \mathbf{v}^j}{\Delta a} \text{ for } j = 1, \dots, J.$$

- **Implicit method**: the HJB can be written as

$$\mathbf{B}^{j+1}\mathbf{v}^j = \mathbf{b}^{j+1}, \quad \text{where} \quad \mathbf{B}^{j+1} = \left(\frac{1}{\Delta a} + \rho - \gamma\right)\mathbf{I} - \tilde{\mathbf{\Pi}}^{j+1}$$

$$\mathbf{b}^{j+1} = \mathbf{u}^{j+1} + \frac{1}{\Delta a}\mathbf{v}^{j+1}.$$

which can be solved efficiently for $\mathbf{v}^j$ with sparse matrix routines

Solving the KF equation

- Define $\mathbf{\Pi}^j$ as $\tilde{\mathbf{\Pi}}^j$ witout the correction terms (i.e., $\xi \equiv 1$)
- Recall the discretized, stacked KF equation + adding up for population:

$$\boldsymbol{\phi}^{j+1} = \left(\mathbf{I} - \Delta a \left(\mathbf{\Pi}^j\right)'\right)^{-1} \boldsymbol{\phi}^j$$
$$\frac{1}{A} = \sum_i \phi^j_{i,1} \Delta x + \sum_i \phi^j_{i,2} \Delta x$$

- We've already computed $\tilde{\mathbf{\Pi}}^j$ to get the HJB, just need to correct and transpose
- Just solve directly for $\boldsymbol{\phi}^{j+1}$ at **almost no extra cost!**
    - Iterate forward from $\boldsymbol{\phi}^1 = \underline{\boldsymbol{\phi}}$
- Renormalize $\boldsymbol{\phi}^j$ if needed to ensure it adds to $1/A$

Reminder of algorithm (we just did steps 1 & 2 in depth)

- **Outer loop**: Guess growth rate γ, solve inner loop, update guess $\gamma(\sigma, \phi)$, repeat.
- **Inner loop**: Given γ, adapt the usual **HACT algorithm** (Achdou et al., 2022)

0. Begin with guess $\{\mathbf{X}^0, \mathbf{p}^0\}$. Denote iterations by $\ell = 0, 1, 2, \ldots$.
1. Given $\{\mathbf{X}^\ell, \mathbf{p}^\ell\}$, solve detrended HJB w/ finite difference method + calculate policy functions $\sigma_n^\ell(a, x)$ and $\mu_n^\ell(a, x)$.
2. Given $\sigma_n^\ell(a, x)$ and $\mu_n^\ell(a, x)$, solve KF for $\phi_n^\ell(a, x)$ w/ finite difference method.
3. Given $\phi_n^\ell(a, x)$, compute vibrancies and housing prices:

$$\tilde{X}_n^\ell = \iint x\, \phi_n^\ell(a, x)\, dx\, da, \qquad \tilde{p}_n^\ell = p_n \left(\iint \phi_n^\ell(a, x)\, dx\, da \right)^\theta.$$

4. If $\{\tilde{\mathbf{X}}^\ell, \tilde{\mathbf{p}}^\ell\}$ close enough to $\{\mathbf{X}^\ell, \mathbf{p}^\ell\}$, **stop**. Else, construct $\{\mathbf{X}^{\ell+1}, \mathbf{p}^{\ell+1}\}$ as a linear combination of previous guess and computed values, then return to step 1.

◂ Quantification

Quantification: Select cities and set congestion elasticities, $\{\mathcal{N}, \theta_n\}$

Spatial scope: 378 MSAs, but smaller are grouped together

- all with 2010 Census pop. > 2mil represented individually $\rightarrow$ **30 MSAs**
  (New York, Los Angeles, Chicago, . . . , Cleveland, Kansas City)
- below, group by 500K $\rightarrow$ **4 additional groups** (*finer partition in progress*)
- groups contain **copies** $\rightarrow$ correct geography & pop. scale, miss within-group variation
- **why?** accord with later regressions from Duranton and Puga (2022)

Quantification: Select cities and set congestion elasticities, $\{\mathcal{N}, \theta_n\}$

Spatial scope: 378 MSAs, but smaller are grouped together

- all with 2010 Census pop. > 2mil represented individually $\rightarrow$ **30 MSAs**
  (New York, Los Angeles, Chicago, ..., Cleveland, Kansas City)
- below, group by 500K $\rightarrow$ **4 additional groups** (*finer partition in progress*)
- groups contain **copies** $\rightarrow$ correct geography & pop. scale, miss within-group variation
- **why?** accord with later regressions from Duranton and Puga (2022)

Congestion elasticity mapped to housing supply (Saiz, 2010)

- housing supply elasticity $= f$(land availability, land use regulations)
- for groups: use pop.-weighted mean

Quantification: Determine migration params., $\{\lambda, \epsilon, \tau_{ni}\}$, using ACS data

2011–15 ACS Migration files count moves within/across MSAs **cross-tabbed** by age

- 5-year average of 1-year migration events, where we can see...
  1. % that didn't move, $1 - \lambda(a)$
  2. % that moved within same MSA, $\mu_{nn}(a)$
  3. % that moved to any other given MSA, $\mu_{ni}(a)$

Quantification: Determine migration params., $\{\lambda, \epsilon, \tau_{ni}\}$, using ACS data

2011–15 ACS Migration files count moves within/across MSAs **cross-tabbed** by age

- 5-year average of 1-year migration events, where we can see...
  1. % that didn't move, $1 - \lambda(a)$
  2. % that moved within same MSA, $\mu_{nn}(a)$
  3. % that moved to any other given MSA, $\mu_{ni}(a)$

So, do the following:

- set $\lambda(a)$ to **match fraction that move by age group** (*decreasing* from 26.6% $\searrow$ 7.9% because marriage, family size, home ownership)
- set $\epsilon = 3$ from Diamond (2016) [w.r.t. real wages at decadal frequency]
- invert bilateral costs from flows using **Head-Ries index**:

$$\mu_{ni}(a,x) = \frac{\tau_{ni}^{-\epsilon} V_i(a,x)^{\epsilon}}{\sum_k \tau_{nk}^{-\epsilon} V_k(a,x)^{\epsilon}} \implies \frac{\bar{\mu}_{ni}\bar{\mu}_{in}}{\bar{\mu}_{nn}\bar{\mu}_{ii}} = \frac{\tau_{ni}^{-\epsilon}\tau_{in}^{-\epsilon}}{\tau_{nn}^{-\epsilon}\tau_{ii}^{-\epsilon}}$$

Quantification: Determine migration params., $\{\lambda, \epsilon, \tau_{ni}\}$, using ACS data

2011–15 ACS Migration files count moves within/across MSAs **cross-tabbed** by age

- 5-year average of 1-year migration events, where we can see...
  1. % that didn't move, $1 - \lambda(a)$
  2. % that moved within same MSA, $\mu_{nn}(a)$
  3. % that moved to any other given MSA, $\mu_{ni}(a)$

So, do the following:

- set $\lambda(a)$ to **match fraction that move by age group**
  (*decreasing* from 26.6% $\searrow$ 7.9% because marriage, family size, home ownership)
- set $\epsilon = 3$ from Diamond (2016) [w.r.t. real wages at decadal frequency]
- invert bilateral costs from flows using **Head-Ries index**:

$$\mu_{ni}(a, x) = \frac{\tau_{ni}^{-\epsilon} V_i(a, x)^{\epsilon}}{\sum_k \tau_{nk}^{-\epsilon} V_k(a, x)^{\epsilon}} \implies \frac{\bar{\mu}_{ni}\bar{\mu}_{in}}{\bar{\mu}_{nn}\bar{\mu}_{ii}} = \frac{\tau_{ni}^{-\epsilon}\tau_{in}^{-\epsilon}}{\tau_{nn}^{-\epsilon}\tau_{ii}^{-\epsilon}}$$

Targeting mobility by age and avg. bilateral flows, not $\mu_{ni}(a, x)$

Quantification: Set human capital investment params., $\{A, \rho, \beta, \eta, \underline{\phi}_n\}$

Key idea: Worker's investment problem nests Ben-Porath (1967) model

$\rightarrow$ **calibrate to previous structural estimates** that used U.S. data

(Heckman, Lochner, and Taber, 1998; Browning, Hansen, and Heckman, 1999; Huggett, Ventura, and Yaron, 2006)

Quantification: Set human capital investment params., $\{A, \rho, \beta, \eta, \underline{\phi}_n\}$

Key idea: Worker's investment problem nests Ben-Porath (1967) model
$\rightarrow$ **calibrate to previous structural estimates** that used U.S. data
(Heckman, Lochner, and Taber, 1998; Browning, Hansen, and Heckman, 1999; Huggett, Ventura, and Yaron, 2006)

Patterns matched:

- concentrate skill acquisition when young; steeper earnings profile if more schooling
- concavity of the cross-sectional earnings distribution across ages
- trends in mean earnings and earnings dispersion & skewness as the typical cohort ages

Quantification: Set human capital investment params., $\{A, \rho, \beta, \eta, \underline{\phi}_n\}$

Key idea: Worker's investment problem nests Ben-Porath (1967) model
$\rightarrow$ **calibrate to previous structural estimates** that used U.S. data
(Heckman, Lochner, and Taber, 1998; Browning, Hansen, and Heckman, 1999; Huggett, Ventura, and Yaron, 2006)

Patterns matched:

- concentrate skill acquisition when young; steeper earnings profile if more schooling
- concavity of the cross-sectional earnings distribution across ages
- trends in mean earnings and earnings dispersion & skewness as the typical cohort ages

Calibration:

- worker's horizon: $A=40$ (age 20–59) and $\rho=0.10$ (discount & IES)
- investment elasticities: $\beta=0.8$ and $\eta=0.7$
- initial human capital distribution, $\underline{\phi}_n$
    - *shape*: log-normal with coefficient of variation 0.468
    - *mean*: varies by HS ($x\approx 10$) vs. COL ($x\approx 13$) $\rightarrow$ weight by 2011–15 ACS college share
    - *mass*: match share of 15-19 year olds in 2010 ACS 1-year sample

Quantification: Minimum distance estimator, $\{\alpha, \zeta, x_{\text{scale}}, B_n, T_n, p_n, \bar{\kappa}\}$

Identify $\{\alpha, \zeta, x_{\text{scale}}\}$ by matching **wage panel regressions** from Duranton and Puga (2022)

$$\ln y_{nt}^j = \boldsymbol{a_n} + a_j + a_t + \sum_n \boldsymbol{b_n} e_{nt}^j + \mathbf{C}_t^j \mathbf{b} + \varepsilon_{nt}^j$$

Find that $\hat{\boldsymbol{a}}_{\boldsymbol{n}}$ and $\hat{\boldsymbol{b}}_{\boldsymbol{n}}$ are generally **increasing** in city size

Quantification: Minimum distance estimator, $\{\alpha, \zeta, x_{\text{scale}}, B_n, T_n, p_n, \bar{\kappa}\}$

Identify $\{\alpha, \zeta, x_{\text{scale}}\}$ by matching **wage panel regressions** from Duranton and Puga (2022)

$$\ln y^j_{nt} = a_n + a_j + a_t + \sum_n b_n e^j_{nt} + \mathbf{C}^j_t \mathbf{b} + \varepsilon^j_{nt}$$

Find that $\hat{a}_n$ and $\hat{b}_n$ are generally **increasing** in city size

1. Differential value of experience: pin down ζ

$$1.0114 = \underbrace{\frac{\hat{b}_{\text{5mil}}}{\hat{b}_{\text{2mil}}}}_{\text{estimate}} = \underbrace{\left(\frac{X_{\text{5mil}}}{X_{\text{2mil}}}\right)^{1-\beta}}_{\text{model}} = \left(\frac{L_{\text{5mil}}^{\frac{1}{\zeta}} \bar{x}_{\text{5mil},\zeta}}{L_{\text{2mil}}^{\frac{1}{\zeta}} \bar{x}_{\text{2mil},\zeta}}\right)^{1-\beta}$$

2. IV of static city FE on city size: pin down α (can match directly without MDE)

$$\hat{a}_n = \alpha \ln L_n + \varepsilon_n$$

3. IV of medium-run city effect on city size: pin down x_{scale}

$$\hat{a}_n + \hat{b}_n \bar{e} = (\alpha + \varsigma) \ln L_n + \varepsilon_n$$

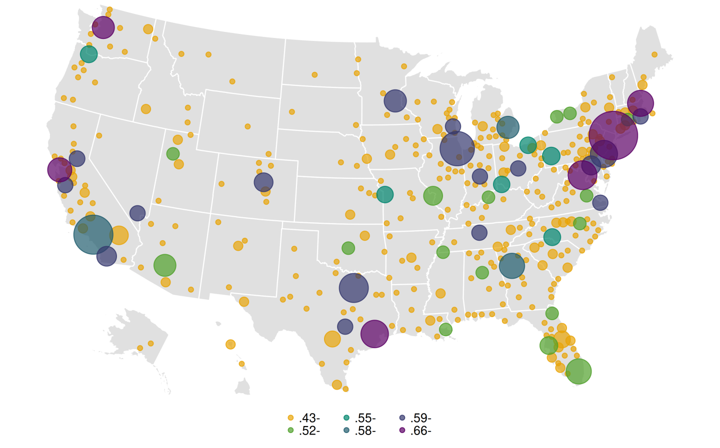
Quantification: Minimum distance estimator, $\{\alpha, \zeta, x_{\text{scale}}, B_n, T_n, p_n, \bar{\kappa}\}$

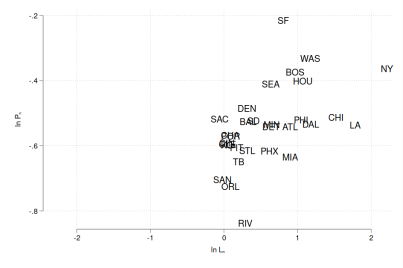
Remaining scales $\{B_n, T_n, p_n, \bar{\kappa}\}$ estimated to minimize distance between model and data for:

- total **employment** in each city per the 2013 BEA Regional Accounts
- the total **wage bill** in each city per the 2013 BEA Regional Accounts
- **constant local expenditure shares** across all cities (Diamond, 2016)
- a **2%** annual growth rate

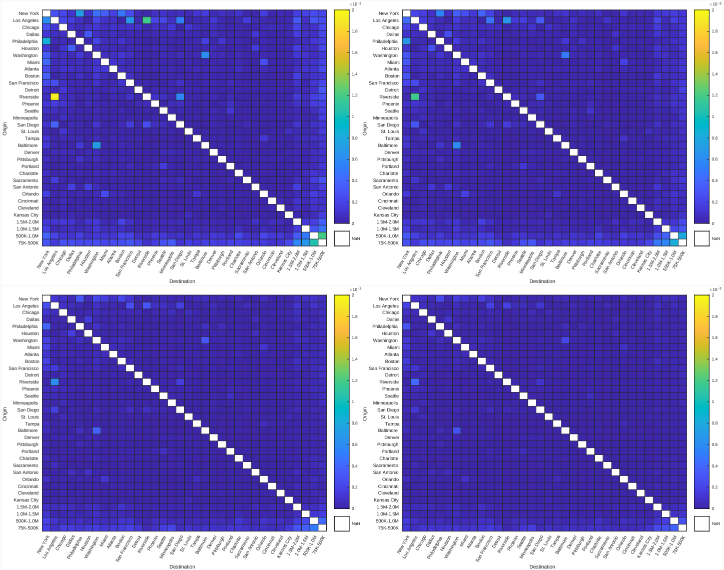
Able to **match exactly** even though cannot invert the model (solving ϕ nonparametrically)

City-level aggregates: Urban cost, $P_n = p_n L_n^{\theta_n}$





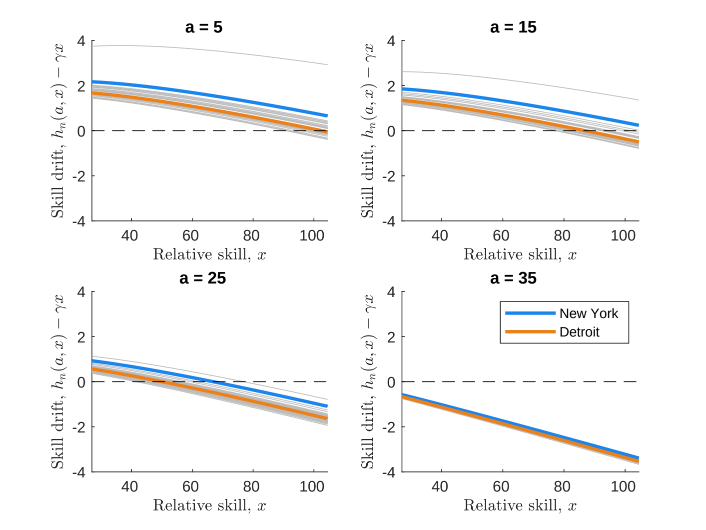
Optimal annual flows, $\lambda(a)\mu_{ni}(a,x)L_n(a)$



$$\mu_{ni}(a,x) = \frac{\tau_{ni}^{-\epsilon} v_i(a,x)^{\epsilon}}{\sum_k \tau_{nk}^{-\epsilon} v_k(a,x)^{\epsilon}}$$

- x (not shown): slight lean to **most vibrant cities**
  - supermodularity $\rightarrow$ PAM (high x with high X_n)
- a: less mobility over time
  - always strong home bias (whited out)
  - when old, stop learning
  - just trade-off income vs. urban cost, both vs. τ_{ni}

Optimal skill *drift*, $h_n(a,x) - \gamma x$



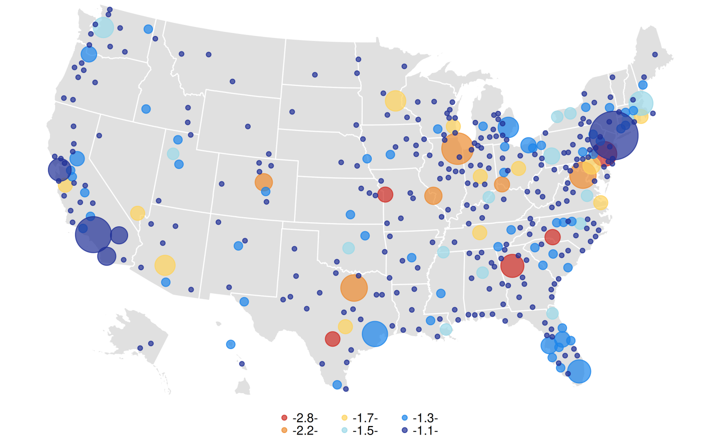
$$d_n(a,x) \equiv \kappa[\sigma_n(a,x)]x^{\beta} X_n^{1-\beta} - \gamma x$$

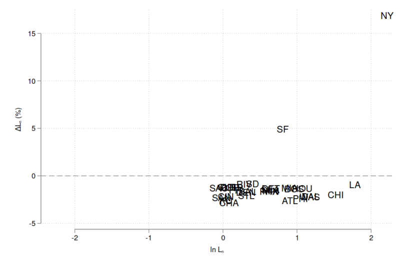
- x: decline w/ relative skill
- a: decline w/ age (zero at A)
- n: inherits from h_n

- $\arg_x d_n(a,x) = 0$ is a **sink**
- density $\phi_n(a,x)$ has finite support if $d_n(a,x)$ has **single-crossing property** of zero in x for all n

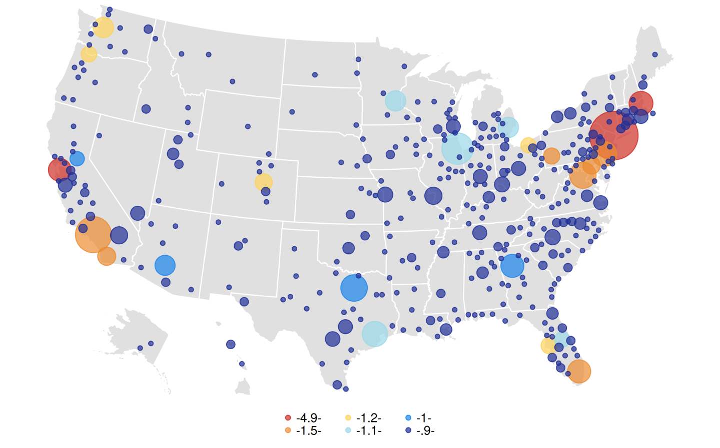
  $\rightarrow$ don't need fat tail

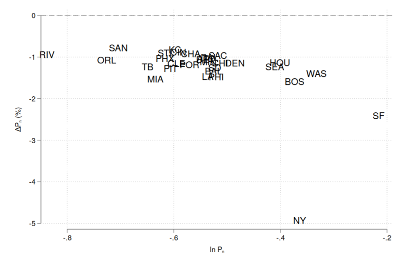
City-level aggregates: ΔL_n (%)



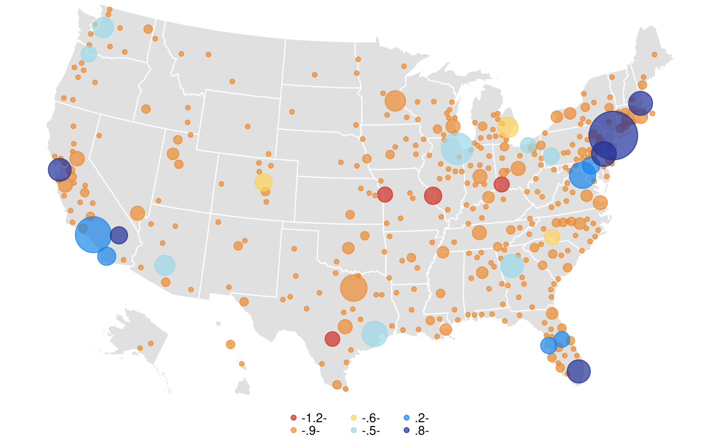


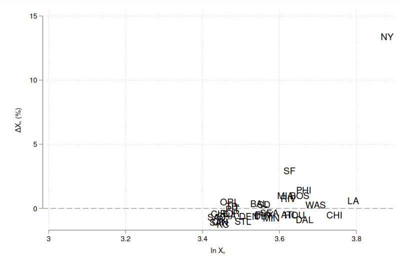
City-level aggregates: ΔP_n (%)



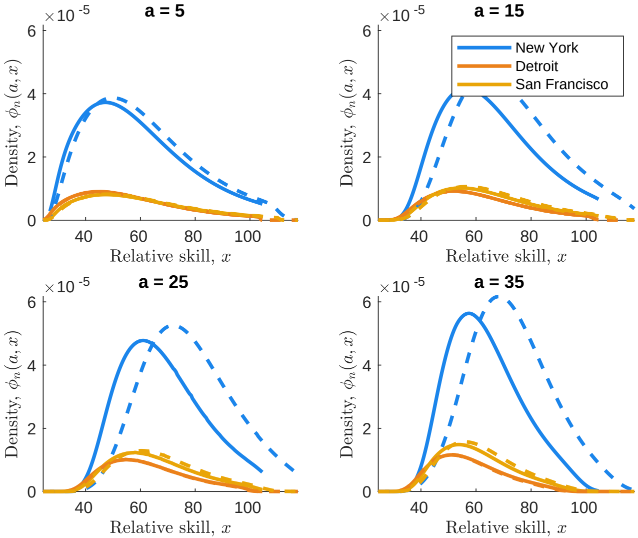


City-level aggregates: ΔX_n (%)

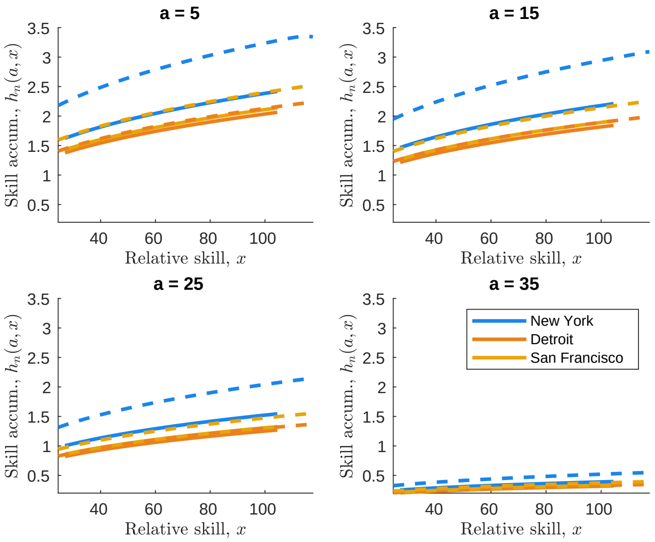




City-level distributions: NY and SF get more skilled, others change little



A new channel for spatial policy: *Produce*, not just *attract*, skill



References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2022. "Income and wealth distribution in macroeconomics: A continuous-time approach." *Review of Economic Studies* 89 (1):45–86.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf. 2015. "The economics of density: Evidence from the Berlin Wall." *Econometrica* 83 (6):2127–2189.

Baum-Snow, Nathaniel and Ronni Pavan. 2012. "Understanding the city size wage gap." *Review of Economic Studies* 79 (1):88–127.

Ben-Porath, Yoram. 1967. "The production of human capital and the life cycle of earnings." *Journal of Political Economy* 75 (4):352–365.

Breinlich, Holger, Gianmarco I. P. Ottaviano, and Jonathan R. W. Temple. 2014. "Regional growth and regional decline." In *Handbook of Economic Growth*, vol. 2, edited by Philippe Aghion and Steven N. Durlauf, chap. 4. Elsevier, 683–779.

Browning, Martin, Lars Peter Hansen, and James J. Heckman. 1999. "Micro data and general equilibrium models." In *Handbook of Macroeconomics*, vol. 1, chap. 8. Elsevier, 543–633.

Buera, Francisco J. and Robert E. Lucas. 2018. "Idea flows and economic growth." *Annual Review of Economics* 10 (1):315–345.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro. 2019. "Trade and labor market dynamics: General equilibrium analysis of the China trade shock." *Econometrica* 87 (3):741–835.

De la Roca, Jorge and Diego Puga. 2017. "Learning by working in big cities." *Review of Economic Studies* 84 (1):106–142.

Desmet, Klaus, Dávid Kristián Nagy, and Esteban Rossi-Hansberg. 2018. "The geography of development." *Journal of Political Economy* 126 (3):903–983.

Diamond, Rebecca. 2016. "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000." *American Economic Review* 106 (3):479–524.

Duranton, Gilles and Diego Puga. 2015. "Urban land use." In *Handbook of Regional and Urban Economics*, vol. 5, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, chap. 8. Elsevier, 467–560.

———. 2022. "Urban growth and its aggregate implications." National Bureau of Economic Research, Working Paper 26591.

Fajgelbaum, Pablo D. and Cecile Gaubert. 2020. "Optimal spatial policies, geography, and sorting." *Quarterly Journal of Economics* 135 (2):959–1036.

Gabriel, Antonio and Robert E. Lucas, Jr. 2019. "Idea flows, economic growth and human capital."

Glaeser, Edward L. 2008. *Cities, agglomeration, and spatial equilibrium*. The Lindahl Lectures. London: Oxford University Press.

———. 2011. *Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier*. New York: Penguin.

Glaeser, Edward L. and David C. Maré. 2001. "Cities and skills." *Journal of Labor Economics* 19 (2):316–342.

Heckman, James J. 1976. "A life-cycle model of earnings, learning, and consumption." *Journal of Political Economy* 84 (4):S11–S44.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents." *Review of Economic Dynamics* 1 (1):1–58.

Hsieh, Chang-Tai and Enrico Moretti. 2019. "Housing constraints and spatial misallocation." *American Economic Journal: Macroeconomics* 11 (2):1–39.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2006. "Human capital and earnings distribution dynamics." *Journal of Monetary Economics* 53 (2):265–290.

Jacobs, Jane. 1969. *The economy of cities*. New York: Vintage Books.

Kennan, John and James R. Walker. 2011. "The effect of expected income on individual migration decisions." *Econometrica* 79 (1):211–251.

Lucas, Robert E., Jr. 1988. "On the mechanics of economic development." *Journal of Monetary Economics* 22 (1):3–42.

Lucas, Robert E., Jr. 2009. "Ideas and growth." *Economica* 76 (301):1–19.

Lucas, Robert E., Jr and Benjamin Moll. 2014. "Knowledge growth and the allocation of time." *Journal of Political Economy* 122 (1):1–50.

Rosen, Sherwin. 1976. "A theory of life earnings." *Journal of Political Economy* 84 (4):S45–S67.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Felipe Schwartzman. 2021. "Cognitive hubs and spatial redistribution." National Bureau of Economic Research, Working Paper 26267.

Saiz, Albert. 2010. "The geographic determinants of housing supply." *Quarterly Journal of Economics* 125 (3):1253–1296.

Uzawa, Hirofumi. 1965. "Optimum technical change in an aggregative model of economic growth." *International Economic Review* 6 (1):18–31.

Wang, Zhi. 2016. "Wage growth, ability sorting, and location choice at labor-force entry: New evidence from US Census data." *Journal of Urban Economics* 96 (1):112–120.