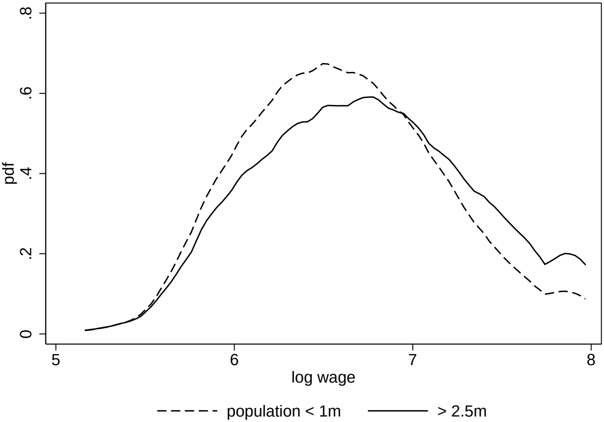
Introduction

- Sinatra effect: Large cities seem to attract talent
- Bankers on Wall Street, actors on Broadway, movie stars in Hollywood, comedians in Las Vegas, software engineers in San Francisco, geologists in Houston's oil business,...
- Well established fact: wages (and output) in large cities are on average higher than in smaller ones
- Is it due to spatial sorting? High skilled citizens locate in large cities, low skilled in small cities?
- But, little evidence of sorting (Baum-Snow & Pavan 2010)

The urban wage premium



$\Delta log = 0.13$ ($t = 27$, $p = 0.0000$)

Between smallest (pop.$\approx$ 160,000) and largest cities: 25%

The lower tail

- We claim that cities also attract the least skilled workers
- Eminem effect
- Despite our priors: Detroit and New York City are very similar to each other
- Large contingent of South Bronx (NY) Newark (NJ), and many talented workers in Bloomfield Hills (Detroit)

Objective

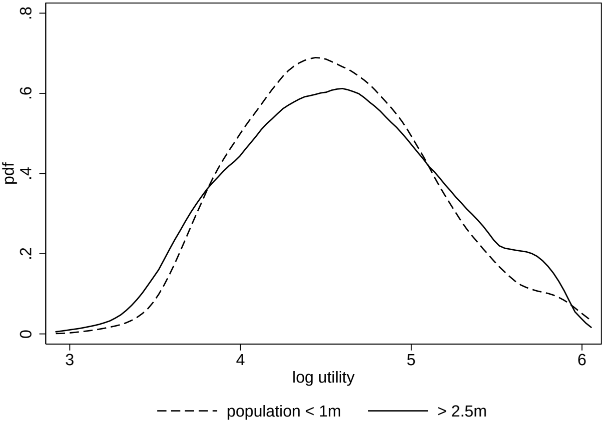
- What we want to know
  - How does the *skill* distribution change across city sizes?
  - What are the mechanisms behind these differences?

Objective

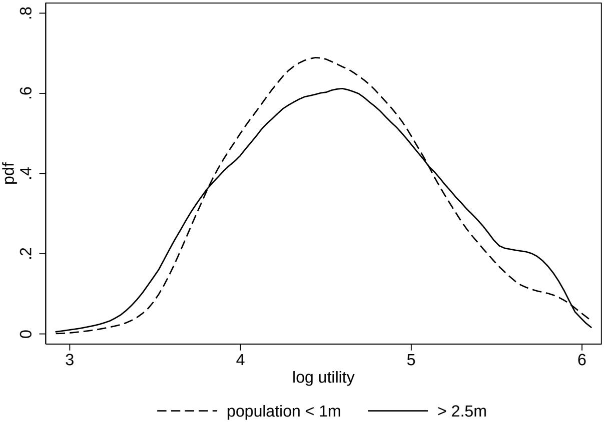
- What we want to know
  - How does the *skill* distribution change across city sizes?
  - What are the mechanisms behind these differences?

- What we do
  - We propose a simple theory of city choice by workers with heterogenous skills.
  - We use the model to study the (unobservable) distribution of skills across cities.

What we find



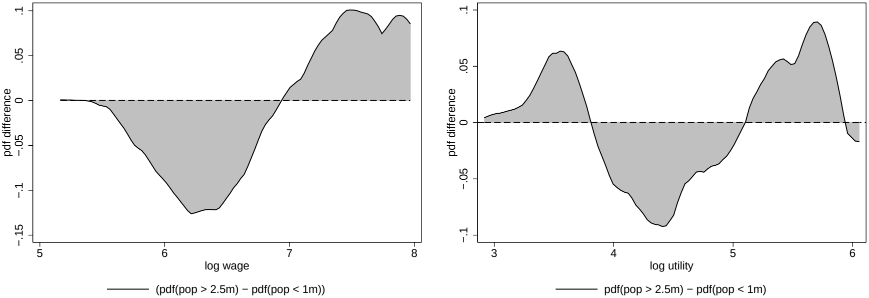
What we find



Spatial Sorting:

1. high and low skilled in large cities
2. medium skilled in small cities

Wage vs. Skill distribution



The model

- J locations (cities) $j \in \mathcal{J} = \{1, ..., J\}$
- Fixed amount of land (housing) H_j

Citizens

- Citizens (workers) with heterogenous skills
  - skill type $i \in \mathcal{I} = \{1, ..., I\}$, measure M_i in economy
  - skill y_i increasing in i
- Preferences are homothetic:

$$u(c_{ij}, h_{ij}) = c_{ij}^{1-\alpha} h_{ij}^{\alpha}$$

c_{ij}: consumption of skill type i in city j
h_{ij}: housing of skill type i in city j
$\alpha \in [0, 1]$: budget share of housing

- Workers are perfectly mobile. Utility equalization across cities:

$$u(c_{ij}, h_{ij}) = u(c_{ij'}, h_{ij'})$$

for any two cities j and j' and for all skill types i.

Technology

- Cities differ exogenously in TFP A_j
- Representative firm in j produces

$$A_j \sum_{i=1}^{I} (m_{ij})^{\gamma_i} y_i^{\beta}$$

m_{ij}: employment level of skill type i in firm in city j
$\gamma_i \in [0,1]$: related to but $\neq$ input elasticity of substitution

- Varying elasticity of substitution (VES): γ_i varies by type i
- Fixed cost of market entry kp_j. Zero profit condition:

$$A_j \sum_{i=1}^{I} (m_{ij})^{\gamma_i} y_i^{\beta} - \sum_{i=1}^{I} w_{ij} m_{ij} - kp_j = 0.$$

w_{ij}: wage for skill type i in city j

Market clearing

- Housing market clears in each city j

$$N_j(k + \sum_{i=1}^{I} h_{ij} m_{ij}) = H_j$$

N_j: measure of firms in j
H_j: fixed housing stock in j

- City population: $S_j = \sum_{i=1}^{I} m_{ij} N_j$

- Labour market clears for each skill type i

$$\sum_{j=1}^{J} N_j m_{ij} = M_i$$

M_i: country-wide measure of skill type i

Citizen's problem

- Workers maximize utility

$$\max_{\{c_{ij},h_{ij}\}} u(c_{ij},h_{ij}) = c_{ij}^{1-\alpha} h_{ij}^{\alpha}$$
$$\text{s.t. } c_{ij} + p_j h_{ij} \leq w_{ij}$$

- Optimal consumption is

$$c_{ij}^{\star} = (1-\alpha) w_{ij}$$
$$h_{ij}^{\star} = \alpha \frac{w_{ij}}{p_j}$$

- Indirect utility function

$$U_i = \alpha^{\alpha} (1-\alpha)^{1-\alpha} \frac{w_{ij}}{p_j^{\alpha}}$$

Firm's problem

- A firm's problem is given by

$$\max_{m_{ij},\forall i} A_j \sum_{i=1}^{I} (m_{ij})^{\gamma_i} y_i^{\beta} - \sum_{i=1}^{I} w_{ij} m_{ij} - k p_j$$

- The first-order condition is for all i

$$\gamma_i A_j (m_{ij})^{\gamma_i - 1} y_i^{\beta} = w_{ij}$$

Assumption 1: Productivity-skill monotonicity

For a given economy, there exists a critical $\beta^\star$ such that for every $\beta > \beta^\star$ productivity is increasing in skill i in every city j.

- *There is a one-to-one relation between skills y_i and equilibrium utility U_i which can be expressed as*

$$U_i = \alpha^\alpha (1-\alpha)^{1-\alpha} \frac{w_{ij}}{p_j^\alpha}$$

Two-city version

- Assume two cities $j = 1, 2$
- From firm's FOC and labour market clearing we derive:

$$\frac{w_{i1}}{w_{i2}} = \left\{ \frac{\frac{A_1}{A_2}\left(\frac{w_1^i}{\gamma_i A_1 y_i^{\beta}}\right)^{\frac{1}{\gamma_i - 1}}}{\frac{M_i}{N_2} - \frac{N_1}{N_2}\left(\frac{w_1^i}{\gamma_i A_1 y_i^{\beta}}\right)^{\frac{1}{\gamma_i - 1}}} \right\}^{\gamma_i - 1}$$

- From utility equalization we derive

$$\frac{w_{i1}}{w_{i2}} = \left(\frac{p_1}{p_2}\right)^{\alpha}$$

Two-city version

- Housing market clearing leads to four equilibrium conditions

$$\sum_{i=1}^{I}\left(\frac{\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_i-1}}}{1+\frac{N_1}{N_2}\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_i-1}}}\frac{M_i}{N_2}\right)^{\gamma_i}\left(1-\left(1-\alpha\right)\gamma_i\right)y_i^{\beta} = \frac{p_1}{A_1}\frac{H_1}{N_1} \quad (1)$$

$$\sum_{i=1}^{I}\left(\frac{1}{1+\frac{N_1}{N_2}\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_i-1}}}\frac{M_i}{N_2}\right)^{\gamma_i}\left(1-\left(1-\alpha\right)\gamma_i\right)y_i^{\beta} = \frac{p_2}{A_2}\frac{H_2}{N_2} \quad (2)$$

$$\sum_{i=1}^{N}\left(\frac{\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_i-1}}}{1+\frac{N_1}{N_2}\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_i-1}}}\frac{M_i}{N_2}\right)^{\gamma_i}\left(1-\gamma_i\right)y_i^{\beta} = \frac{k}{A_1}p_1 \quad (3)$$

$$\sum_{i=1}^{N}\left(\frac{1}{1+\frac{N_1}{N_2}\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_i-1}}}\frac{M_i}{N_2}\right)^{\gamma_i}\left(1-\gamma_i\right)y_i^{\beta} = \frac{k}{A_2}p_2 \quad (4)$$

Main results

Theorem 1: CES technology

If $\gamma_i = \gamma$ for all i, then the skill distribution across cities is identical

Main results

Theorem 1: CES technology

If $\gamma_i = \gamma$ for all i, then the skill distribution across cities is identical

- Homotheticity technology $\Rightarrow$ demand diff. skills proportional

Main results

Proposition 1: City Size and TFP

More productive cities are larger. Let $A_1 > A_2$ *and* $H_1 = H_2$*, then* $S_1 > S_2$.

Main results

Proposition 1: City Size and TFP

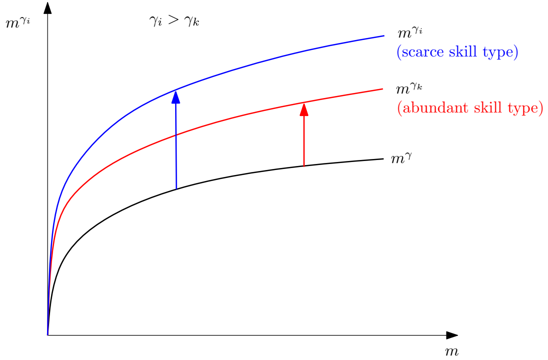
More productive cities are larger. Let $A_1 > A_2$ and $H_1 = H_2$, then $S_1 > S_2$.

- $H_1 = H_2$ is not necessary, but extremely scarce supply of land can increase p_1 so much that $S_1 < S_2$

Assumption

Assumption 2:

γ_i is decreasing in supply of skill i, M_i.



Main results

Theorem 2: Fat Tails

Consider a symmetric, uni-modal skill distribution economy-wide. Then under Assumption 1 and $H_1 = H_2$, the skill distribution in larger cities has fatter tails.

Fat tails

- We seek to check whether really both the lower and the upper tail of larger cities are fatter than in smaller ones (Theorem 2).
- Skills are at best partially observable.

⇒ We use theory to estimate skills as

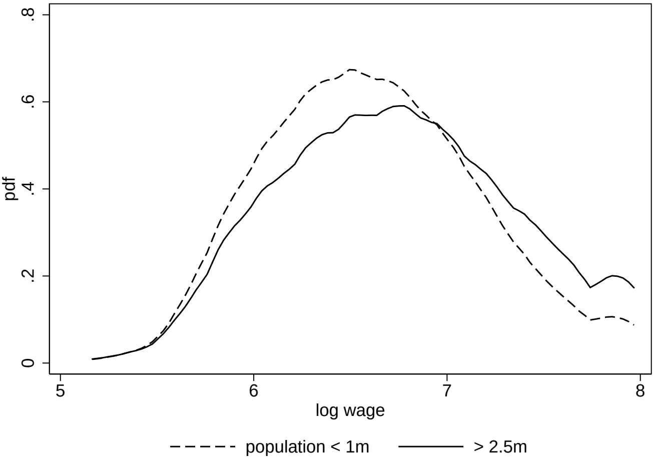
$$U_i = \alpha^{\alpha} (1-\alpha)^{1-\alpha} \frac{w_{ij}}{p_j^{\alpha}}$$

- Need to observe:
  - wage distribution w_{ij} by city
  - housing price level p_j
  - budget share of housing α

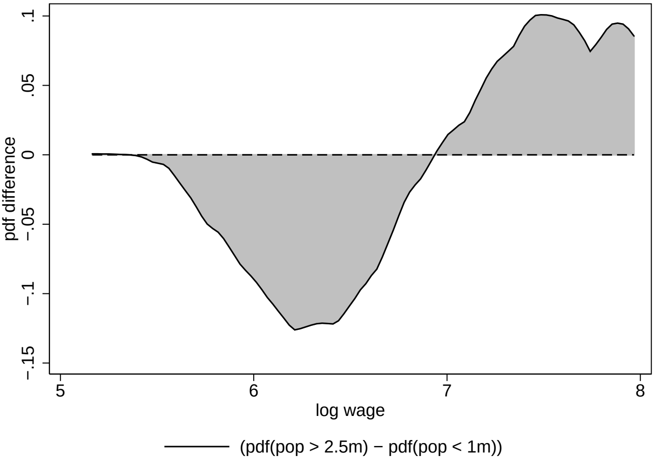
Wages

- *Data Source*: Current population survey (CPS) 2009.
- Weekly earnings for 102,577 workers in 257 U.S. metropolitan areas with 2009 population above 100,000.
- Most comprehensive definition of metro area: core based statistical area (CBSA).
- Need to take care of top-coding ($150,000).

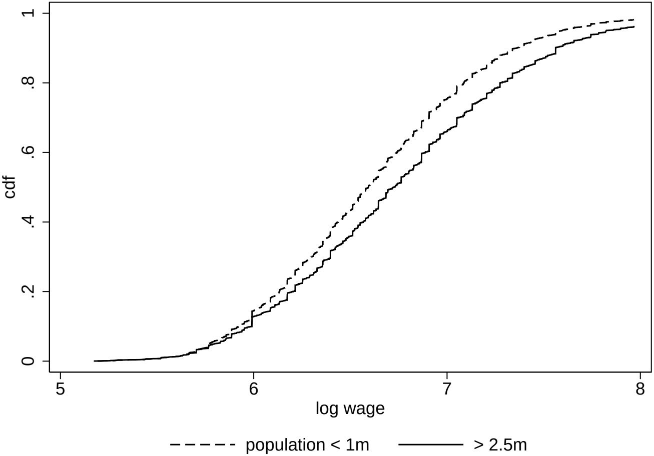
Wages and city size



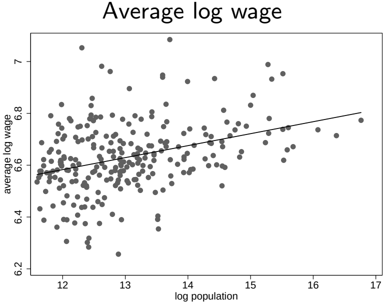
Wages and city size



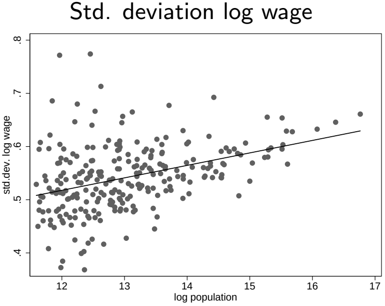
Wages and city size



Wages and city size



slope = .046
(t=7.07, p=0.000)



slope = 0.023
(t=8.03, p=0.000)

Housing prices

- *Data Source*: Decennial Census of Housing 2000. 5% Public Use Microsample (PUMS).
- Monthly rent for 3,274,198 housing units and assessed housing value for 7,680,728 units in 1887 areas (PUMA).
- Aggregated to 533 core based statistical area (CBSA).
- Need to control for quality differences in housing.

⇒ Hedonic regression on observed housing characteristics

- $e^{cityfixedeffect}$ is housing price index
- Need to adjust for 2000-2009 growth

⇒ Repeat Sales price growth index from Federal Housing Finance Agency (FHFA).

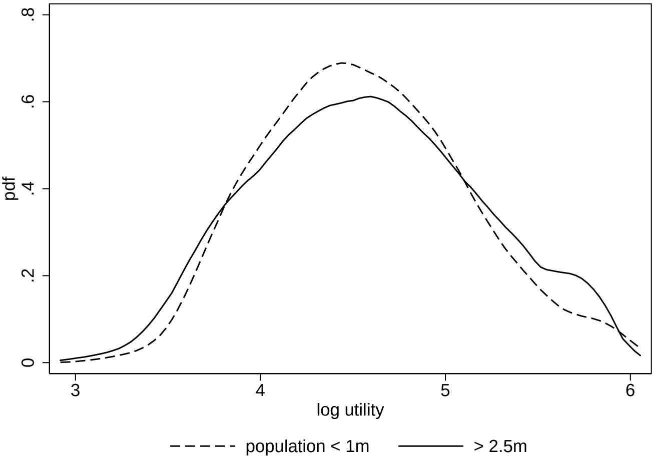
Skills

⇒ We use theory to estimate skills as

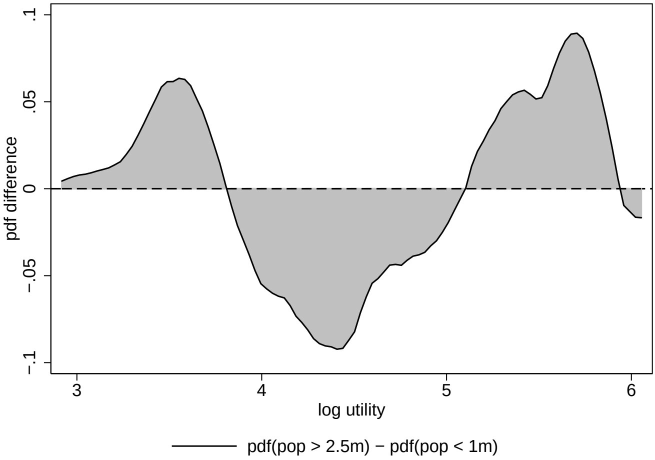
$$U_i = \alpha^\alpha (1-\alpha)^{1-\alpha} \frac{w_{ij}}{p_j^\alpha}$$

- $\hat{\alpha} = 0.24$ from Davis and Ortalo-Magné (RED 2010)

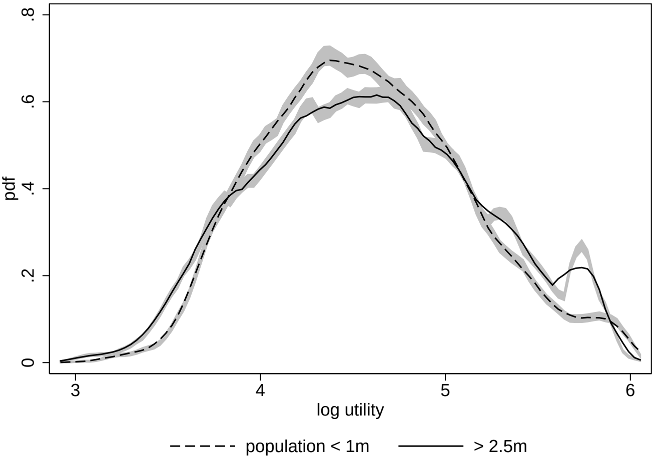
Skills and city size



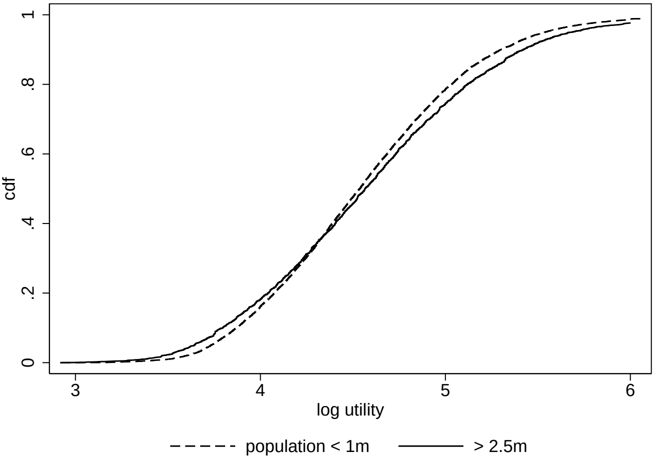
Skills and city size



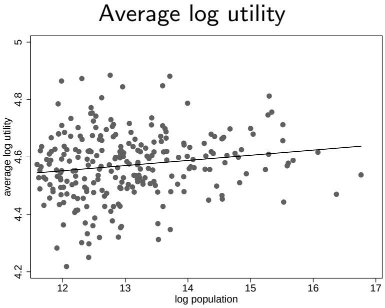
Skills and city size



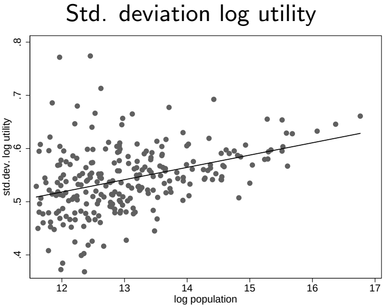
Skills and city size



Skills and city size



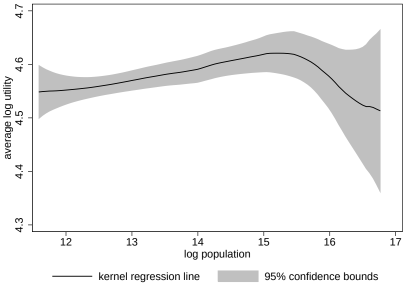
slope = .018
(t=2.90, p=0.004)



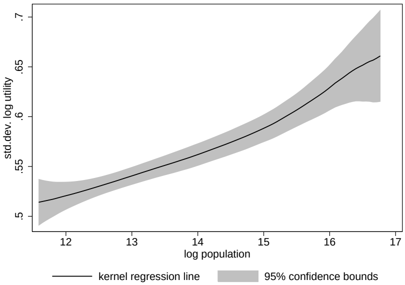
slope = 0.023
(t=7.79, p=0.000)

Skills and city size

Average log utility



Std. deviation log utility

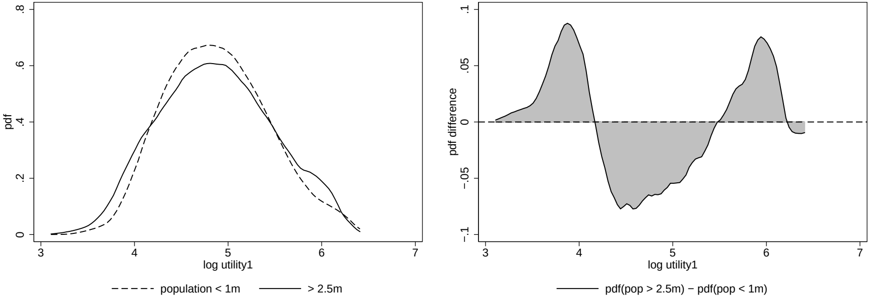


Discussion and Extensions

- Use different cut-offs: very robust
- Use alternative house price indices: very robust
- Consider all local prices: robust
- Compare *observed* skills across city sizes: robust
- Conditonal *unobserved* skills across city sizes: robust
- Explore alternative explanations:
  - Migration: consistent
  - Labor Market Frictions: competing
  - Specialization/Division of Labor: competing

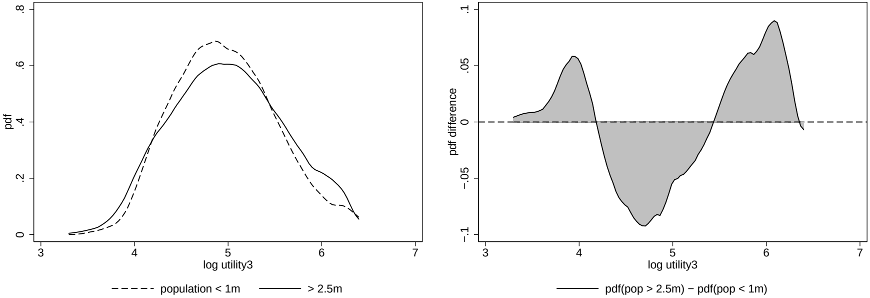
Alternative Housing Prices Measures

National Association of Realtors, median house value 2009



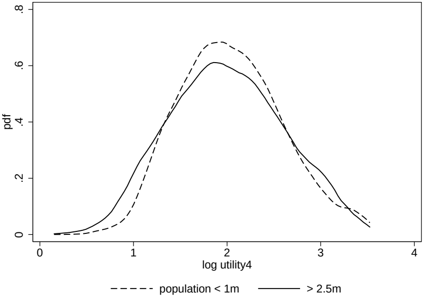
Alternative Housing Prices Measures

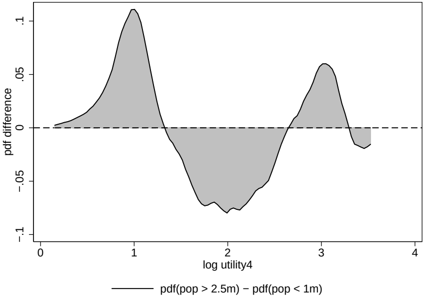
ACCRA, local housing price index



All Local Prices

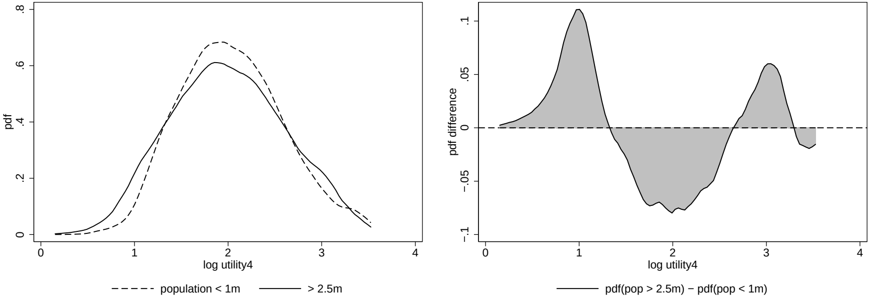
ACCRA, local price indices for 6 broad categories





All Local Prices

ACCRA, local price indices for 6 broad categories



Local price indices for grocery items (sausage), housing (rent), utilities (phone call), transportation (gasoline), health care (Lipitor) and services (haircut).

ACCRA data is problematic:

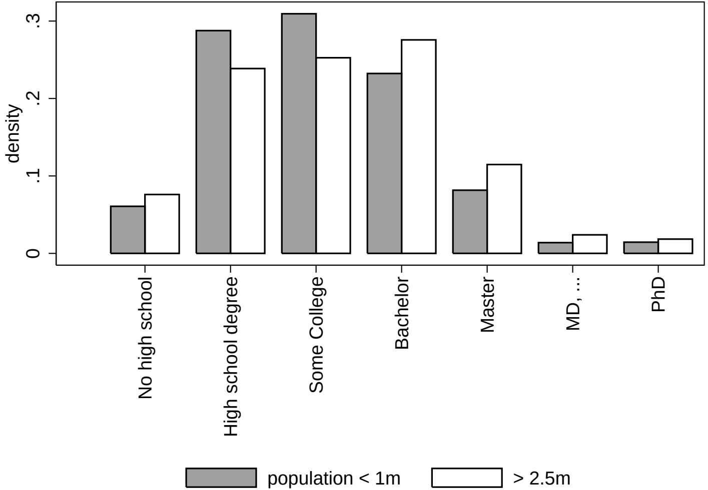
- very noisy from volunteer data collection
- based on arbitrary 60! items
- Does not correct (enough) for quality differences: a "haircut" in NYC is not the same as a haircut in Springfield

→ Likely to overstate price differentials

⇒ We see above figure as *upper bound*

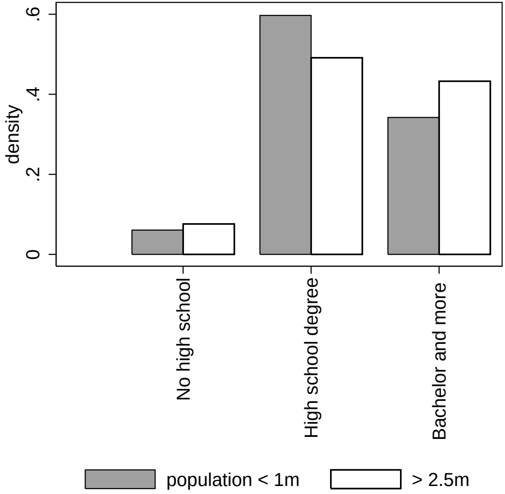
Observable Measures of Skill

Education and city size



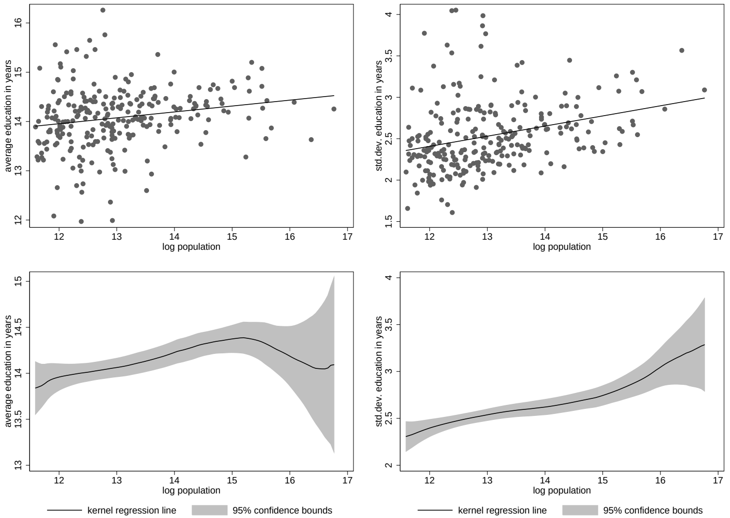
Observable Measures of Skill

Three Categories



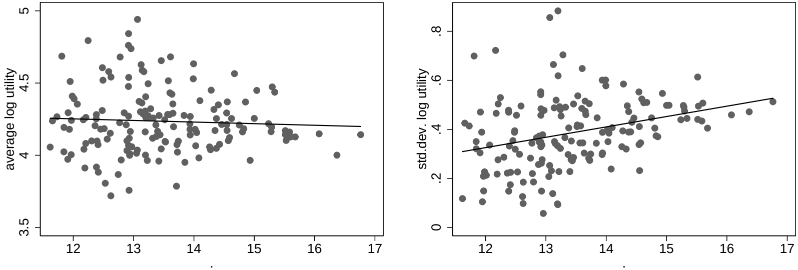
Observable Measures of Skill

Education in (usual) years



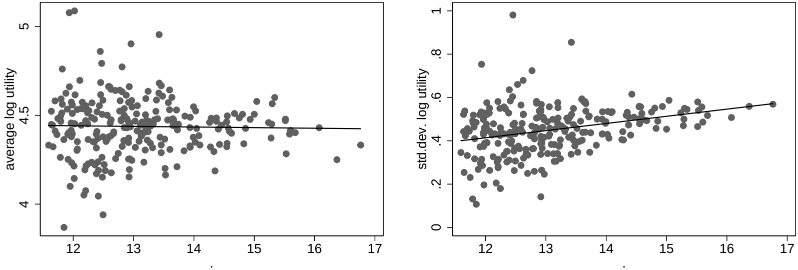
Conditional unobserved skills

Education level 1: No high school



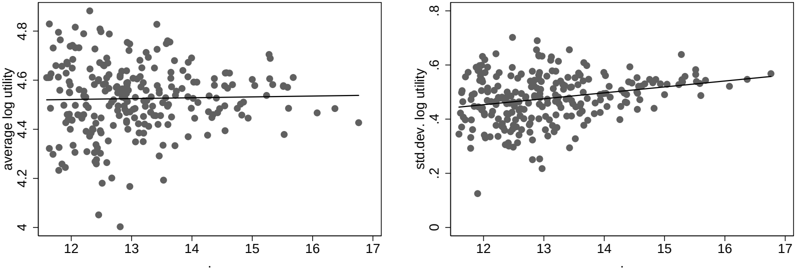
Conditional unobserved skills

Education level 2: High school

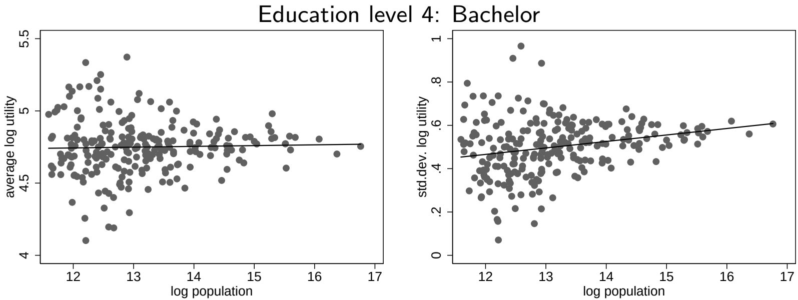


Conditional unobserved skills

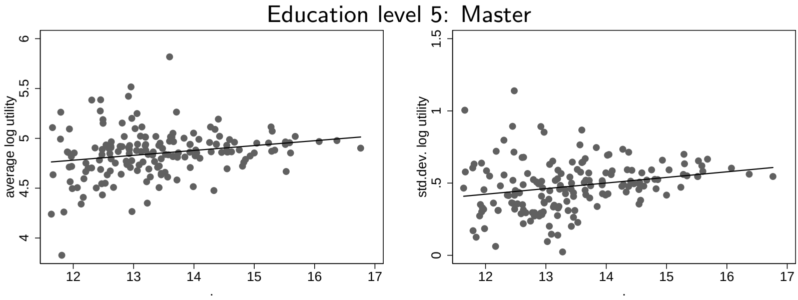
Education level 3: Some college



Conditional unobserved skills

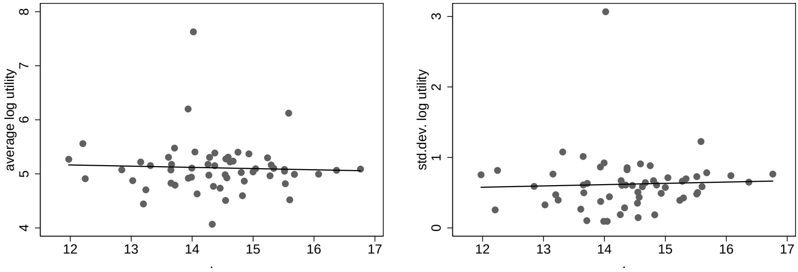


Conditional unobserved skills



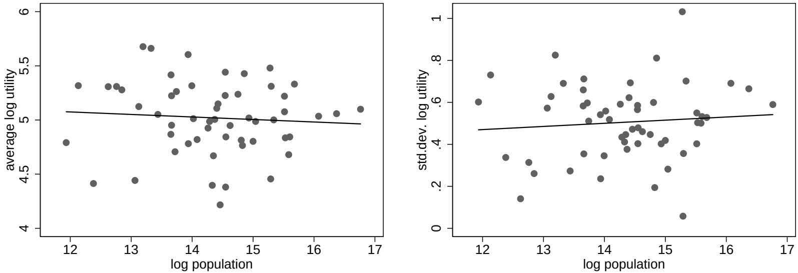
Conditional unobserved skills

Education level 6: MD, JD, ...



Conditional unobserved skills

Education level 7: PhD

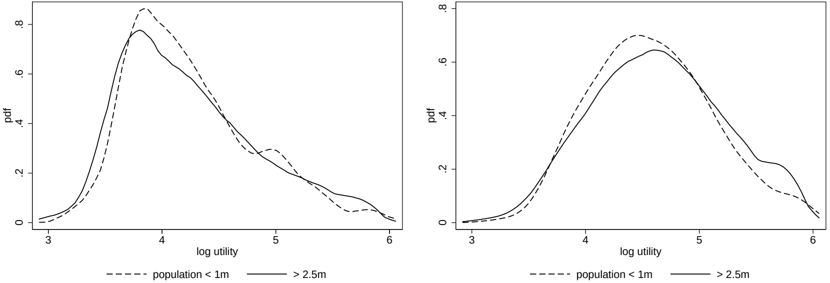


The role of migration

- Immigrants are disproportionately located in large cities. For example, because of spillovers from immigrant networks.
- Evidence from CPS:
  - foreign born are more likely to live in large cities (76% vs 55% for natives)
  - foreign born are a larger fraction of workforce in large cities (12% vs 5% in small cities).
- Could the observed fatter lower tail be because immigrants are disproportionately low skilled?
  1. It does not matter for theory: arbitrage of marginal type between small and large cities
  2. Evidence, no. Depending on the total numbers, some natives will also live in large cities or some immigrants in small cities.

The role of migration

Fat Tails



- Fat tails for *both* A. Foreign born; B. Natives
- Also high skilled, foreign born move to large cities

Conclusion

- Well established fact: Urban wage premium
- We show that the *average* wage premium can almost fully be explained as compensation for higher housing prices
- Is there evidence of spatial sorting?
- The really interesting attribute lies in the second moments: We document higher variance of skills in large cities
- We propose a (simple) theory of city choice by heterogenous workers and show that this empirical regularity
  - is not possible under standard CES technology
  - can be rationalized by a VES production technology
- Methodological: given non-monotonic relative demand, cannot capture sorting by partitioning skills into two classes

Productive cities:
Sorting, selection and agglomeration

Kristian Behrens

Gilles Duranton

Frédéric Robert-Nicoud

Objectives

- Provide a unified model of an urban system consistent with recently uncovered stylised facts about cities

Plan of the talk

- Stylised facts
- Model
- Sorting equilibrium
- Adding numbers to the theory
- Other equilibria (in progress)

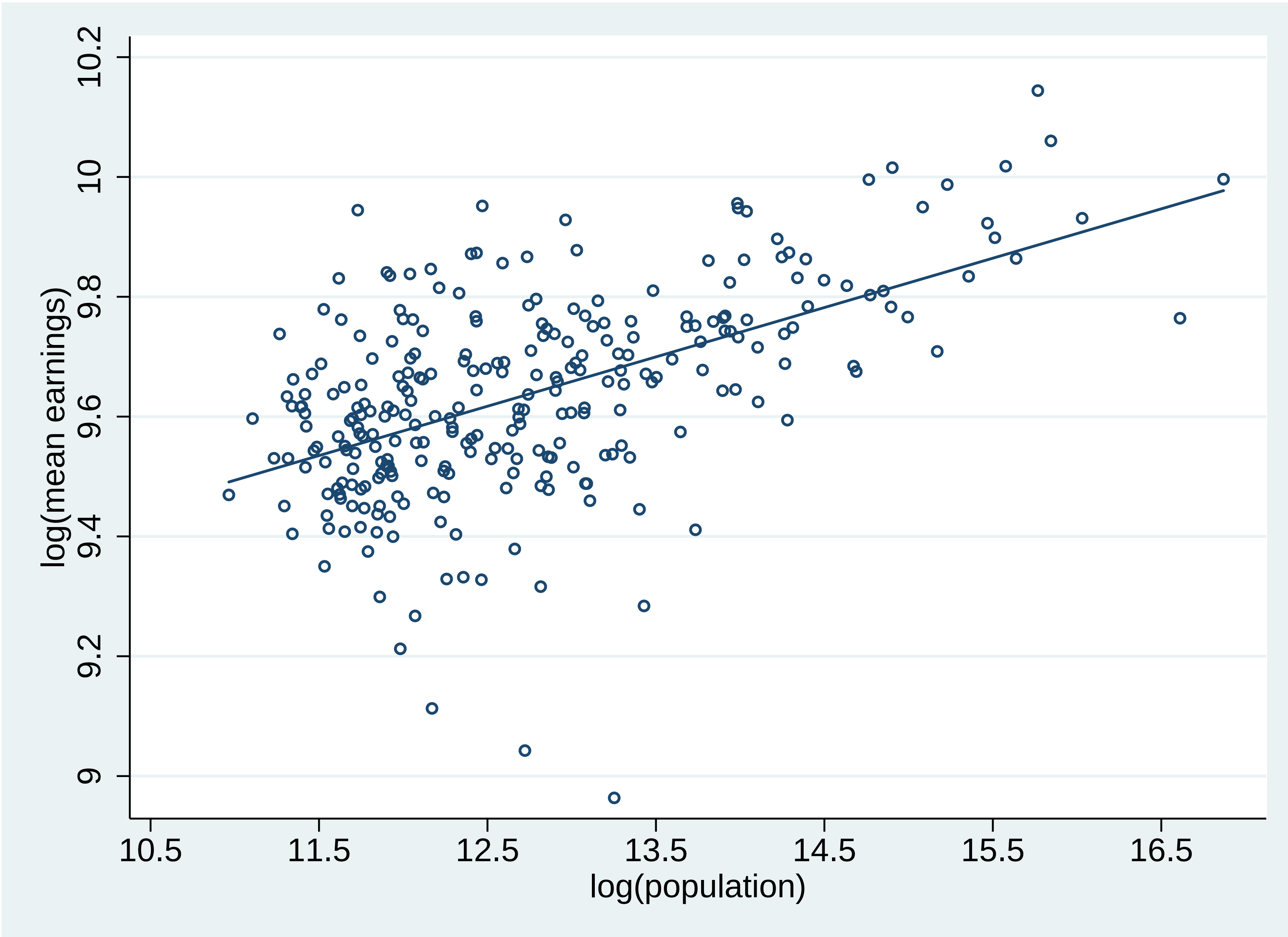


Figure 1: The city size - productivity relationship

Stylised fact 1: Agglomeration economies

- Positive association between various measures of scale and productivity
- Causal? IV evidence: Ciccone and Hall (1996); Combes, Duranton, Gobillon, and Roux (2010)
- Causal? Quasi-experimental evidence: Greenstone, Hornbeck, and Moretti (2008)
- Causal? Robust to ability sorting (Combes, Duranton, and Gobillon, 2008) and market selection (Combes, Duranton, Gobillon, Puga, and Roux, 2009)
- Needed to justify the existence of cities
- Input-output linkages as a key channel (Holmes, 1999; Amiti and Cameron, 2007; Overman and Puga, 2009; and Ellison, Glaeser, and Kerr, 2010).

Stylised fact 2: Ability sorting across cities

- Direct evidence of mild sorting:
  - Population elasticity of the share of college graduates: 6.8% (US, 2000)
  - Well documented by the literature (Berry and Glaeser, 2005; Bacolod, Blum, and Strange, 2009; Lee, 2010)
- But appears to account for up to half the size productivity relationship (Combes, Duranton, and Gobillon, 2008; Baum-Snow and Pavan, 2009a)
- Higher relative rewards for skilled workers in larger cities (Wheeler, 2001; Bacolod, Blum, and Strange, 2009; Glaeser and Resseger, 2010)
- Selective migration (Dahl, 2002).

Stylised fact 3: Market selection

- Could explain the size productivity relationship (Sinatra, 1979)
- ‘Aggregate’ evidence of selection (Bartelsman and Doms, 2000; Foster, Haltiwanger, and Syverson, 2008)
- Higher survival productivity cutoff in larger markets (Syverson, 2004)
- No evidence of tougher selection in larger markets after conditioning out agglomeration and sorting (Combes, Duranton, Gobillon, Puga, and Roux, 2009)
- The share of self-employed is independent from city population in the US

Bottom line: market selection exists but not directly observed in urban data

Stylised fact 4: Size distribution of cities

Zipf's law (Pareto distribution with unitary shape parameter)

- Reasonable approximation of existing size distribution of cities (Gabaix and Ioannides, 2004; Soo 2005)

Related literature

- Little on sorting across (large) locations (Abdel-Rahman and Wang, 1997; Mori and Turrini, 2005; Baldwin and Okubo, 2006; Nocke, 2006)
- More closely related: Behrens and Robert-Nicoud (2009) and Davis (2009)
- Zipf's law typically explained by random growth models (Gabaix, 1999; Eeckhout, 2004; Duranton, 2007; Rossi-Hansberg and Wright, 2007)
- Two static models (Hsu, 2008; Lee and Li, 2009)
- System of cities (Henderson, 1974)

Model: Timing

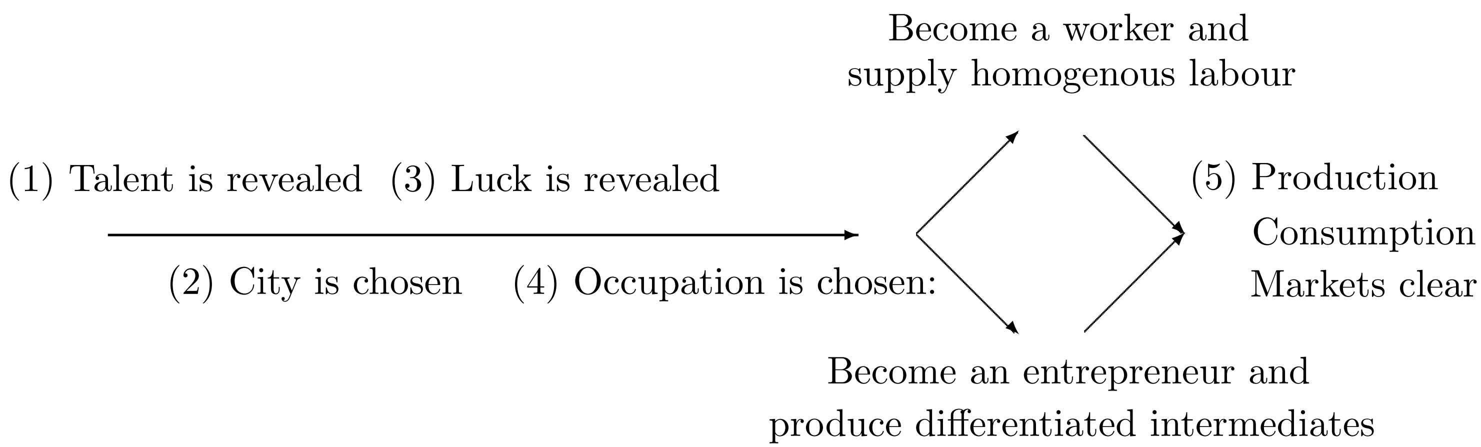


Figure 2: Timing

Talent and luck determine entrepreneurial productivity: $\varphi \equiv t \times s$.

Model: Preferences

Risk-neutral individuals consume one unit of land and a final good.

Model: Technology

Two-step production process

- Aggregate output in city c:

$$Y_c = \left[\int_{\Omega_c^+} x_c(i)^{\frac{1}{1+\varepsilon}} \mathrm{d}i \right]^{1+\varepsilon}$$

- Freely tradable across cities (numéraire)
- Local intermediates:

$$x_c(i) = \varphi(i) l_c(i),$$

  with $\varphi(i) = t(i) \times s(i)$

In short: Dixit and Stiglitz (1977) - Ethier (1982) - Melitz (2003)

Model: Short-run equilibrium

Selection: Individuals become either workers or entrepreneurs

Solve for occupations, prices and quantities in each city. Each individual takes as given

- Own productivity,
- Cumulative productivity distribution $F_c(\cdot)$
- Population size L_c

Profit maximisation implies:

$$p(i) = (1+\varepsilon)\frac{w}{\varphi(i)}$$

and

$$\pi(i) = \frac{\varepsilon}{1+\varepsilon}\, p(i)\, x(i) = \frac{\varepsilon}{1+\varepsilon}\, Y \left[\frac{\varphi(i)}{\Phi}\right]^{\frac{1}{\varepsilon}}$$

with

$$\Phi \equiv \left[\int_{\Omega^+} \varphi(j)^{\frac{1}{\varepsilon}}\, \mathrm{d}j\right]^{\varepsilon}$$

a measure of aggregate productivity in the city.

Entrepreneur if $\pi(i) > w$. Cutoff productivity:

$$\underline{\varphi} \equiv \Phi\left(\frac{1+\varepsilon}{\varepsilon}\;\frac{w}{Y}\right)^{\varepsilon}. \tag{1}$$

Demand and supply on the labour market:

$$L^S = F(\underline{\varphi})L\,.$$

$$L^D = L\int_{\underline{\varphi}}^{\sup\Omega^+} l(\varphi)\mathrm{d}F(\varphi) = \frac{1}{1+\varepsilon}\;\frac{Y}{w},$$

Yield equilibrium:

$$Y = (1+\varepsilon)F(\underline{\varphi})Lw\,, \tag{2}$$

Profit maximisation by final producers:

$$w = \frac{1}{1+\varepsilon} \mathit{\Phi}, \tag{3}$$

Finally, aggregate profit:

$$\mathit{\Pi} = \frac{\varepsilon}{1+\varepsilon} \mathit{Y} = \varepsilon F(\underline{\varphi}) w L \tag{4}$$

Existence and uniqueness of short-run equilibrium

Proposition 1 *Given population, L, and its productivity distribution, $F(\cdot)$, the equilibrium in a city exists, is unique, and characterised by* (1) *to* (4).

Fixed point of:

$$F(\underline{\varphi}) = \frac{1}{\varepsilon} \int_{\underline{\varphi}}^{\sup \Omega^{+}} \left(\frac{\varphi}{\underline{\varphi}} \right)^{\frac{1}{\varepsilon}} \mathrm{d}F(\varphi).$$

Short-run equilibrium properties

Proposition 2 *Given the productivity distribution, $F(\cdot)$, larger cities have higher aggregate productivity, per-capita income, and wages than smaller cities. The productivity cutoff for selection does not depend on city size.*

Aggregate city income as

$$Y = F(\underline{\varphi})\Phi L = F(\underline{\varphi})\left[\int_{\underline{\varphi}}^{\sup \Omega^+} \varphi^{\frac{1}{\varepsilon}} \mathrm{d}F(\varphi)\right]^{\varepsilon} L^{1+\varepsilon}.$$

The productivity cutoff for selection does not depend on city size!

A useful knife edge?

Model: Urban structure

- One-dimension monocentric city with commuting costs $t(x) = \tau x^\gamma$ where $\gamma > \varepsilon$
- With unitary lot sizes, residential equilibrium: $\tau x^\gamma + r(x) = \tau(L/2)^\gamma + r(L/2) = \tau(L/2)^\gamma$
- Total land rent is redistributed

$$\text{TLR} = 2\int_0^{L/2} r(x)\mathrm{d}x = \frac{2\tau\gamma}{\gamma+1}\left(\frac{L}{2}\right)^{\gamma+1} .$$

- Utility is income minus average commuting cost

$$\text{ACC} = \frac{2}{L}\int_0^{L/2} t(x)\mathrm{d}x = \frac{\text{TLR}}{\gamma L} = \theta L^\gamma$$

with $\theta \equiv 2^{-\gamma}\tau(\gamma+1)^{-1}$

Long-run equilibrium

- City population and talents are endogenous: $L_c \equiv \int_{\underline{t}}^{\bar{t}} L_c(t) \mathrm{d}t$
- Individuals make a location choice based on their talent to maximise $\mathbb{E} V_c(t)$
- No individual wants to deviate to another city given the location choices of all other individuals
- Cumulative distribution of luck is the same in all cities

The complementarity between talent and city size

Individual t's expected indirect utility in c:

$$\begin{aligned} \mathbb{E}V_c(t) &= \int_{\underline{s}}^{\overline{s}} \max\{w_c, \pi_c(ts)\} \mathrm{d}G_s(s) - \theta L_c^{\gamma} \\ &= w_c G_s(\underline{\varphi}_c / t) + w_c \left(\frac{t}{\underline{\varphi}_c} \right)^{\frac{1}{\varepsilon}} \int_{\underline{\varphi}_c / t}^{\overline{s}} s^{\frac{1}{\varepsilon}} \mathrm{d}G_s(s) - \theta L_c^{\gamma} \end{aligned}$$

Proposition 3 *More talented individuals benefit disproportionately from being located in larger cities:*

$$\frac{\partial^2 \mathbb{E}V_c(t)}{\partial t \partial L_c} \geq 0$$

Important caveat: Proposition 3 is for a given $\underline{\varphi}_c$

Equilibrium with talent-homogenous cities

For now, focus on equilibrium with a single type of talent t_c in each city

Lemma 1 *In talent-homogeneous cities, the productivity cutoff is proportional to talent: there exist* $\underline{s} < S < \overline{s}$ *and* $\phi \in (0,1)$ *such that* $\underline{\varphi}_c = St_c$ *and* $G_s(S) = \phi$ *for all* $c \in C$*, with*

$$\phi \equiv G_s(S) = \frac{1}{\varepsilon} \int_S^{\overline{s}} \left(\frac{s}{S} \right)^{\frac{1}{\varepsilon}} \mathrm{d}G_s(s).$$

Two properties:

- Sorting induces selection
- Conditional on sorting, no differences in selection across cities

Constructing the equilibrium with talent-homogenous cities

Define the matching function $\mu : T \to C$ (talents into cities). An equilibrium is a correspondence $c = \mu(t)$ such that, for all $t \in T$ and for all c,$c' \in$ C:

$$\mu(t) = \left\{ c : \mathbb{E}V_c(t) \geq \mathbb{E}V_{c'}(t),\ \forall c' \in C \right\}.$$

Convex problem: order cities so that $t_c = t(c)$ and $L_c = L(c)$ are continuous functions of c.

The first-order condition is

$$\left.\frac{\partial \mathbb{E} V_c(t)}{\partial L_c}\right|_{t=t_c} \mathrm{d}L_c + \left.\frac{\partial \mathbb{E} V_c(t)}{\partial t_c}\right|_{t=t_c} \mathrm{d}t_c = 0. \tag{5}$$

- Individual t's utility is maximised in the city where all individuals have the same talent as hers
- The first term of (5) defines a constrained social optimum
- The second term is positive
- Individuals have an incentive to move to a city less talented than socially optimal
- More talented individuals will want to be in more talented cities only if they are larger
- Cities are too large in equilibrium

In equilibrium the FOC holds for $t = t_c$ and city size is a function of talent: $L_c = L(t_c)$. Equation (5) yields differential equation

$$\gamma\theta L(t_c)^{\varepsilon}\left[\xi t_c \frac{L'(t_c)}{L(t_c)} - L(t_c)^{\gamma-\varepsilon}\frac{L'(t_c)}{L(t_c)} + \delta\right] = 0,$$

where

$$\xi \equiv \frac{S}{\gamma\theta}\left(\varepsilon\phi\right)^{1+\varepsilon} \quad \text{and} \quad \delta \equiv \frac{\xi}{1+\varepsilon} > 0.$$

Optimal city size under perfect sorting

Proposition 4 *Talent-homogenous cities of optimal size are such that:*

$$L^o(t_c) = (\xi t_c)^{\frac{1}{\gamma-\varepsilon}}.$$

Optimal size increases with talent, t_c, and agglomeration economies, ϵ, decreases with urban costs, θ and γ, is such that the Henry George Theorem holds: $\mathrm{TLR}_c = \varepsilon Y_c$, *and leads to expected indirect utility:* $\mathbb{E}V^0(t_c) = \frac{\theta}{\varepsilon}(\gamma - \varepsilon)[L^o(t_c)]^\gamma$.

Equilibrium city size under perfect sorting

Proposition 5 *Any equilibrium with talent-homogenous cities is unique and such that*

$$L^*(t_c) = \left(\frac{1+\gamma}{1+\varepsilon}\right)^{\frac{1}{\gamma-\varepsilon}} L_c^o(t_c).$$

Equilibrium size is too large, increases with talent, t_c*, and agglomeration economies,* ϵ*, decreases with urban costs,* θ *and* γ*, and leads to expected indirect utility:* $\mathbb{E}V^*(t_c) = \frac{\theta}{\varepsilon}\frac{\gamma-\varepsilon}{1+\gamma}[L^*(t_c)]^\gamma$.

SOCs are restrictive

The size distribution of cities

Proposition 6 (Number of cities) *The equilibrium 'number' of cities is proportional to population size* Λ *and too small relative to the social optimum.*

Proposition 7 (Size distribution of cities) *When talent follows a truncated Pareto distribution with shape parameter* m *over* $[\underline{t}, \bar{t}]$*, the size distribution of talent-homogenous cities is then a truncated Pareto with shape parameter* $1 + (\gamma - \varepsilon)m$ *at both equilibrium and optimum.*

Summary

- Full urban system with endogenous cities
- Agglomeration, selection and sorting interact to explain the urban premium
- Selection has only indirect effects
- Zipf's law can be generated
- Turn now to empirical issues

The city size - productivity relationship

From our model

$$\ln y_c = [\ln(\phi S) + \varepsilon \ln(\varepsilon\phi)] + \ln t_c + \varepsilon \ln L_c$$

and

$$\ln y_c = \ln\left(\frac{1+\varepsilon}{1+\gamma}\,\frac{\gamma\theta}{\varepsilon}\right) + \gamma \ln L_c$$

(recall $t_c = \frac{1+\varepsilon}{(1+\gamma)\xi} L_c^{\gamma-\varepsilon}$)

In the data (US MSAS, 2000):

$$\begin{aligned} \ln y_c &= 9.60 + 0.051 \ln L_c + 0.46 \ln t_c \\ \ln y_c &= 8.59 + 0.082 \ln L_c \end{aligned}$$

Not conditioning out t_c yields an estimate of γ not ϵ

0.082: A reasonable estimate of housing costs?

Using housing rent data

$$\ln r_c = 5.19 + 0.085 \ln L_c$$

Using a more sophisticated housing cost index:

$$\ln h_c = 9.72 + 0.11 \ln L_c$$

Mild sorting, strong effects?

From the model:

$$\ln t_c = \ln \frac{1+\varepsilon}{(1+\gamma)\xi} + (\gamma - \varepsilon) \ln L_c$$

Indirectly from the data:

$$\gamma - \epsilon = 0.082 - 0.051 = 0.031$$

$$\text{or} \quad \gamma - \epsilon = 0.11 - 0.051 = 0.059$$

Directly from the data:

$$\ln t_c = -2.21 + 0.068 \ln L_c.$$

The city oversize debate

From the model and the data:

$$\frac{\widehat{L_c^*}}{L_c^o} = \left(\frac{1+\widehat{\gamma}}{1+\widehat{\varepsilon}}\right)^{\frac{1}{\hat{\gamma}-\hat{\varepsilon}}} = 2.55.$$

(actually $L_c^*/L_c^o \to e$ when ε and $\gamma \to 0$)

Welfare cost of oversize:

$$\begin{aligned}\widehat{\Delta\mathbb{E}V} &\equiv \frac{\mathbb{E}V(\hat{L}^*) - \mathbb{E}V(\hat{L}^o)}{\mathbb{E}V(\hat{L}^o)} \\ &= -1 + \frac{\widehat{\gamma}}{\widehat{\gamma}-\widehat{\varepsilon}}\left(\frac{\widehat{L_c^*}}{L_c^o}\right)^{\widehat{\varepsilon}} - \frac{\widehat{\varepsilon}}{\widehat{\gamma}-\widehat{\varepsilon}}\left(\frac{\widehat{L_c^*}}{L_c^o}\right)^{\widehat{\gamma}} \\ &= -.2\%.\end{aligned}$$

Other equilibria

- Symmetric equilibrium stable under some conditions
- Use simulations to construct equilibria with non-extreme partitions
  - What happens when SOCs are not satisfied for talent homogenous cities
  - ‘Robustness’ of the properties of the extreme partition equilibrium
  - Get interesting results on urban inequalities
  - Get interesting results on productivity distribution (‘dilation’)

Conclusion

- Model of spatial sorting of heterogenous agent in a urban hierarchy
- Key ingredients: Ex ante sorting, ex post selection and agglomeration
- Captures key stylised facts
- Useful to interpret empirical evidence
- ‘Limited spatial arbitrage’
- Way forwards:
  - More work needed on alternative equilibria
  - Make cities less passive