Agglomeration and Sorting

Yujiang Chen and Coen Teulings

University of Utrecht/Cambridge

c.n.teulings@outlook.com

2018

Just one example

- Around 1920, Frits Philips started electric light bulbs factory
- He considered several villages: Helmond, Veghel, and Veldhoven
- In the end, he choose the village of Eindhoven
- Eindhoven: by 1950, 7th city of the Netherlands, by 1970 ranker 5th
- Eindhoven/Philips started a technical university, worldfamous labs
- 1970': dramatic period for Philips, it went almost bankrupt
- Philips' headquarters moved to Amsterdam
- Eindhoven went through a deep trough
- By now: a high tech city, hosting ASML company and many others

Introduction

- Lucas (1988,2001,2009) spillovers, HC, entrepreneurship
- Lucas & Rossi-Hansberg (2002) spatial structure of a city
- Rossi-Hansberg & Wright (2007) distribution across cities
- Gennaioli et.al. (2013)
    - cross section analysis of 1500 regions in 110 countries
    - strong regional sorting of HC within each country
    - high "interregional return to HC": 20-40%
    - interregional spillovers due to HC, not size per se
    - most of the return due to entrepreneurship
- Ahlfeldt et al. (2015) Berlin wall
- Diamond (2016) consumption amenities
- Hsieh & Moretti (2018) housing supply constraints
    - we assume competitive land markets
    - alternative theory of high house prices

Model

- general equilibrium static model of regional sorting
- heterogeneous supply **and** demand for labour
- job assignment / comparative advantage
- non-homothetic, non-Cobb Douglas utility
    - land, amenities, other consumption
    - land & amenities are normal goods
- perfectly elastic interregional labour mobility
- heterogeneous knowledge spillovers
- regions choose between urban and rural structure

Data

- 81 regions in US: 47 states & 34 cities (MSA's)
- labour demand: occupations
- labour supply: years of education et.al.
- house prices
  - Cobb Douglas production of housing services
  - house price reveal land prices
- amenity: January temperature

Conclusions

1. Large interregional wage differentials and returns to HC
2. Simple variables explain 80% interregional wage variation
3. High HC demand activities concentrated in cities
4. Strong agglomeration externalities
5. ... driven by demand for HC (**not** supply!)
6. Demand for HC clusters more strongly than supply
7. ... despite the fact that housing is a normal good
8. Cities must create a lot of amenities to generate this outcome

Menu of the day

1. Some empirical regularities
2. Theoretical model
3. Estimation
4. Counter factual simulation
5. Method generating demand and supply indices

Simple statistical model

$$w = \omega_r + \beta_r h + \varepsilon$$

- - w: log wage of an individual
  - h: Human Capital (HC) index
    - roughly: years of education
  - ε: error term
  - ω_r: regional fixed effect
  - β_r: regional return to HC
  - o: occupational complexity index
    - nation wide mean log wage in occupation
  - H_r and O_r regional means of h and o
  - WLOG: $\mathrm{E}[h] = \mathrm{E}[\omega_r] = \mathrm{E}[o] = \mathrm{E}[1-\beta_r] = 0$
  - hence: $\mathrm{E}[H_r] = \mathrm{E}[O_r] = 0$

Results on intercepts and returns

			Correlation Matrix			
Variable	Mean	S.D.	$\widehat{H}_r$	$\widehat{O}_r$	$\widehat{\omega}_r$	β_r
$\widehat{H}_r$	0.0014	0.0338	1			
$\widehat{O}_r$	0.0010	0.0320	0.7775	1		
$\widehat{\omega}_r$	-0.0645	0.0811	0.6219	0.7909	1	
β_r	0.9872	0.0460	0.1217	0.4673	0.5598	1

Results on intercepts and returns

			Correlation Matrix			
Variable	Mean	S.D.	$\widehat{H}_r$	$\widehat{O}_r$	$\widehat{\omega}_r$	β_r
$\widehat{H}_r$	0.0014	0.0338	1			
$\widehat{O}_r$	0.0010	0.0320	0.7775	1		
$\widehat{\omega}_r$	-0.0645	0.0811	0.6219	0.7909	1	
β_r	0.9872	0.0460	0.1217	0.4673	0.5598	1

Results on intercepts and returns

			Correlation Matrix			
Variable	Mean	S.D.	$\widehat{H}_r$	$\widehat{O}_r$	$\widehat{\omega}_r$	β_r
$\widehat{H}_r$	0.0014	0.0338	1			
$\widehat{O}_r$	0.0010	0.0320	0.7775	1		
$\widehat{\omega}_r$	-0.0645	0.0811	0.6219	0.7909	1	
β_r	0.9872	0.0460	0.1217	0.4673	0.5598	1

Results on intercepts and returns

			Correlation Matrix			
Variable	Mean	S.D.	$\widehat{H}_r$	$\widehat{O}_r$	$\widehat{\omega}_r$	β_r
$\widehat{H}_r$	0.0014	0.0338	1			
$\widehat{O}_r$	0.0010	0.0320	0.7775	1		
$\widehat{\omega}_r$	-0.0645	0.0811	0.6219	0.7909	1	
β_r	0.9872	0.0460	0.1217	0.4673	0.5598	1

Results on intercepts and returns

			Correlation Matrix			
Variable	Mean	S.D.	$\widehat{H}_r$	$\widehat{O}_r$	$\widehat{\omega}_r$	β_r
$\widehat{H}_r$	0.0014	0.0338	1			
$\widehat{O}_r$	0.0010	0.0320	0.7775	1		
$\widehat{\omega}_r$	-0.0645	0.0811	0.6219	0.7909	1	
β_r	0.9872	0.0460	0.1217	0.4673	0.5598	1

Results on intercepts and returns

			Correlation Matrix			
Variable	Mean	S.D.	$\widehat{H}_r$	$\widehat{O}_r$	$\widehat{\omega}_r$	β_r
$\widehat{H}_r$	0.0014	0.0338	1			
$\widehat{O}_r$	0.0010	0.0320	0.7775	1		
$\widehat{\omega}_r$	-0.0645	0.0811	0.6219	0.7909	1	
β_r	0.9872	0.0460	0.1217	0.4673	0.5598	1

First results on agglomeration

	(1)	(2)	(3)	(4)	(5)
VARIABLES			Adj. Avg. Wage ω_r		
Human Capital Index H_r	0.321	0.214	-0.0157	-0.338	-0.517
	(3.07)	(2.74)	(-0.20)	(-2.72)	(-3.00)
Occupation Index O_r				0.580	0.730
				(2.89)	(3.05)
City Dummy			-0.313	-0.309	-0.298
			(-3.08)	(-5.61)	(-4.51)
City x ln Population			0.0265	0.0244	0.0235
			(3.68)	(6.34)	(5.03)
Spatial Lag		0.885	0.755	0.683	0.652
		(8.10)	(7.63)	(8.11)	(7.45)
South Dummy					-0.0226
					(-2.26)
Constant	-0.0666	0.00122	-0.0358	-0.0305	-0.0247
	(-9.42)	(0.13)	(-3.02)	(-3.27)	(-2.82)
R-squared	0.094	0.477	0.683	0.757	0.771
R-MSE	0.0639	0.0489	0.0385	0.0340	0.0332

First results on agglomeration

	(1)	(2)	(3)	(4)	(5)
VARIABLES			Adj. Avg. Wage ω_r		
Human Capital Index H_r	0.321	0.214	-0.0157	-0.338	-0.517
	(3.07)	(2.74)	(-0.20)	(-2.72)	(-3.00)
Occupation Index O_r				0.580	0.730
				(2.89)	(3.05)
City Dummy			-0.313	-0.309	-0.298
			(-3.08)	(-5.61)	(-4.51)
City x ln Population			0.0265	0.0244	0.0235
			(3.68)	(6.34)	(5.03)
Spatial Lag		0.885	0.755	0.683	0.652
		(8.10)	(7.63)	(8.11)	(7.45)
South Dummy					-0.0226
					(-2.26)
Constant	-0.0666	0.00122	-0.0358	-0.0305	-0.0247
	(-9.42)	(0.13)	(-3.02)	(-3.27)	(-2.82)
R-squared	0.094	0.477	0.683	0.757	0.771
R-MSE	0.0639	0.0489	0.0385	0.0340	0.0332

First results on agglomeration

	(1)	(2)	(3)	(4)	(5)
VARIABLES			Adj. Avg. Wage ω_r		
Human Capital Index H_r	0.321	0.214	-0.0157	-0.338	-0.517
	(3.07)	(2.74)	(-0.20)	(-2.72)	(-3.00)
Occupation Index O_r				0.580	0.730
				(2.89)	(3.05)
City Dummy			-0.313	-0.309	-0.298
			(-3.08)	(-5.61)	(-4.51)
City x ln Population			0.0265	0.0244	0.0235
			(3.68)	(6.34)	(5.03)
Spatial Lag		0.885	0.755	0.683	0.652
		(8.10)	(7.63)	(8.11)	(7.45)
South Dummy					-0.0226
					(-2.26)
Constant	-0.0666	0.00122	-0.0358	-0.0305	-0.0247
	(-9.42)	(0.13)	(-3.02)	(-3.27)	(-2.82)
R-squared	0.094	0.477	0.683	0.757	0.771
R-MSE	0.0639	0.0489	0.0385	0.0340	0.0332

First results on agglomeration

	(1)	(2)	(3)	(4)	(5)
VARIABLES			Adj. Avg. Wage ω_r		
Human Capital Index H_r	0.321	0.214	-0.0157	-0.338	-0.517
	(3.07)	(2.74)	(-0.20)	(-2.72)	(-3.00)
Occupation Index O_r				0.580	0.730
				(2.89)	(3.05)
City Dummy			-0.313	-0.309	-0.298
			(-3.08)	(-5.61)	(-4.51)
City x ln Population			0.0265	0.0244	0.0235
			(3.68)	(6.34)	(5.03)
Spatial Lag		0.885	0.755	0.683	0.652
		(8.10)	(7.63)	(8.11)	(7.45)
South Dummy					-0.0226
					(-2.26)
Constant	-0.0666	0.00122	-0.0358	-0.0305	-0.0247
	(-9.42)	(0.13)	(-3.02)	(-3.27)	(-2.82)
R-squared	0.094	0.477	0.683	0.757	0.771
R-MSE	0.0639	0.0489	0.0385	0.0340	0.0332

First results on agglomeration

VARIABLES	(1)	(2)	(3)	(4)	(5)
			Adj. Avg. Wage ω_r		
Human Capital Index H_r	0.321	0.214	-0.0157	-0.338	-0.517
	(3.07)	(2.74)	(-0.20)	(-2.72)	(-3.00)
Occupation Index O_r				0.580	0.730
				(2.89)	(3.05)
City Dummy			-0.313	-0.309	-0.298
			(-3.08)	(-5.61)	(-4.51)
City x ln Population			0.0265	0.0244	0.0235
			(3.68)	(6.34)	(5.03)
Spatial Lag		0.885	0.755	0.683	0.652
		(8.10)	(7.63)	(8.11)	(7.45)
South Dummy					-0.0226
					(-2.26)
Constant	-0.0666	0.00122	-0.0358	-0.0305	-0.0247
	(-9.42)	(0.13)	(-3.02)	(-3.27)	(-2.82)
R-squared	0.094	0.477	0.683	0.757	0.771
R-MSE	0.0639	0.0489	0.0385	0.0340	0.0332

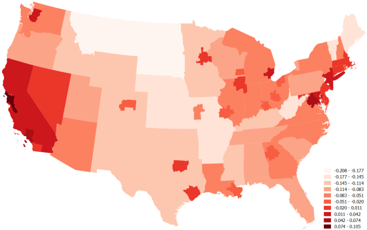
First results on agglomeration

	(1)	(2)	(3)	(4)	(5)
VARIABLES			Adj. Avg. Wage ω_r		
Human Capital Index H_r	0.321	0.214	-0.0157	-0.338	-0.517
	(3.07)	(2.74)	(-0.20)	(-2.72)	(-3.00)
Occupation Index O_r				0.580	0.730
				(2.89)	(3.05)
City Dummy			-0.313	-0.309	-0.298
			(-3.08)	(-5.61)	(-4.51)
City x ln Population			0.0265	0.0244	0.0235
			(3.68)	(6.34)	(5.03)
Spatial Lag		0.885	0.755	0.683	0.652
		(8.10)	(7.63)	(8.11)	(7.45)
South Dummy					-0.0226
					(-2.26)
Constant	-0.0666	0.00122	-0.0358	-0.0305	-0.0247
	(-9.42)	(0.13)	(-3.02)	(-3.27)	(-2.82)
R-squared	0.094	0.477	0.683	0.757	0.771
R-MSE	0.0639	0.0489	0.0385	0.0340	0.0332

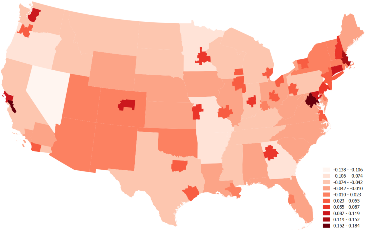
First results on agglomeration

	(1)	(2)	(3)	(4)	(5)
VARIABLES			Adj. Avg. Wage ω_r		
Human Capital Index H_r	0.321	0.214	-0.0157	-0.338	-0.517
	(3.07)	(2.74)	(-0.20)	(-2.72)	(-3.00)
Occupation Index O_r				0.580	0.730
				(2.89)	(3.05)
City Dummy			-0.313	-0.309	-0.298
			(-3.08)	(-5.61)	(-4.51)
City x ln Population			0.0265	0.0244	0.0235
			(3.68)	(6.34)	(5.03)
Spatial Lag		0.885	0.755	0.683	0.652
		(8.10)	(7.63)	(8.11)	(7.45)
South Dummy					-0.0226
					(-2.26)
Constant	-0.0666	0.00122	-0.0358	-0.0305	-0.0247
	(-9.42)	(0.13)	(-3.02)	(-3.27)	(-2.82)
R-squared	0.094	0.477	0.683	0.757	0.771
R-MSE	0.0639	0.0489	0.0385	0.0340	0.0332

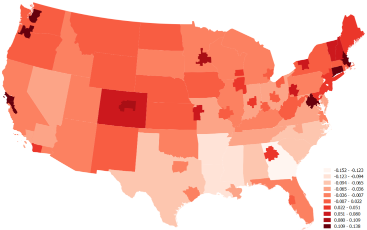
Wages, human capital, and occupations



Average Regional Wage, 1979-2015



Average Occupation Index, 1979-2015



Average Human Capital Index, 1979-2015

Simple model of return to HC

- Return to Human Capital

$$\beta_r = 1 - \gamma \left(h - o \right)$$

- Taking expectations within region r

$$\mathrm{E}[\beta_r | r] = 1 - \gamma \left(H_r - O_r \right)$$

Simple model of return to HC

- Return to Human Capital

$$\beta_r = 1 - \gamma\,(h - o)$$

- Taking expectations within region r

$$\mathrm{E}[\beta_r | r] = 1 - \gamma\,(H_r - O_r)$$

Simple model of return to HC

- Return to Human Capital

$$\beta_r = 1 - \gamma\,(h - o)$$

- Taking expectations within region r

$$\mathrm{E}[\beta_r | r] = 1 - \gamma\,(H_r - O_r)$$

First results on return to HC

VARIABLES	(1) β_r	(2) β_r	(3) β_r	(4) $H_r - O_r$
H_r	-0.442			
	(-4.65)			
O_r	0.717		0.207	-0.118
	(7.25)		(2.13)	(-1.14)
$H_r - O_r$		-0.547		
		(-4.79)		
City Dummy			-0.205	0.0564
			(-2.23)	(0.41)
City x ln Population			0.0162	-0.00471
			(2.55)	(-0.50)
Constant	0.987	0.988	0.976	0.00569
	(239.43)	(221.85)	(141.38)	(0.86)
R-squared	0.366	0.244	0.281	0.081
R-MSE	0.0371	0.0403	0.0398	0.0406

First results on return to HC

VARIABLES	(1) β_r	(2) β_r	(3) β_r	(4) $H_r - O_r$
H_r	-0.442			
	(-4.65)			
O_r	0.717		0.207	-0.118
	(7.25)		(2.13)	(-1.14)
$H_r - O_r$		-0.547		
		(-4.79)		
City Dummy			-0.205	0.0564
			(-2.23)	(0.41)
City x ln Population			0.0162	-0.00471
			(2.55)	(-0.50)
Constant	0.987	0.988	0.976	0.00569
	(239.43)	(221.85)	(141.38)	(0.86)
R-squared	0.366	0.244	0.281	0.081
R-MSE	0.0371	0.0403	0.0398	0.0406

First results on return to HC

VARIABLES	(1) β_r	(2) β_r	(3) β_r	(4) $H_r - O_r$
H_r	-0.442			
	(-4.65)			
O_r	0.717		0.207	-0.118
	(7.25)		(2.13)	(-1.14)
$H_r - O_r$		-0.547		
		(-4.79)		
City Dummy			-0.205	0.0564
			(-2.23)	(0.41)
City x ln Population			0.0162	-0.00471
			(2.55)	(-0.50)
Constant	0.987	0.988	0.976	0.00569
	(239.43)	(221.85)	(141.38)	(0.86)
R-squared	0.366	0.244	0.281	0.081
R-MSE	0.0371	0.0403	0.0398	0.0406

First results on return to HC

VARIABLES	(1) β_r	(2) β_r	(3) β_r	(4) $H_r - O_r$
H_r	-0.442			
	(-4.65)			
O_r	0.717		0.207	-0.118
	(7.25)		(2.13)	(-1.14)
$H_r - O_r$		-0.547		
		(-4.79)		
City Dummy			-0.205	0.0564
			(-2.23)	(0.41)
City x ln Population			0.0162	-0.00471
			(2.55)	(-0.50)
Constant	0.987	0.988	0.976	0.00569
	(239.43)	(221.85)	(141.38)	(0.86)
R-squared	0.366	0.244	0.281	0.081
R-MSE	0.0371	0.0403	0.0398	0.0406

First results on return to HC

VARIABLES	(1) β_r	(2) β_r	(3) β_r	(4) $H_r - O_r$
H_r	-0.442			
	(-4.65)			
O_r	0.717		0.207	-0.118
	(7.25)		(2.13)	(-1.14)
$H_r - O_r$		-0.547		
		(-4.79)		
City Dummy			-0.205	0.0564
			(-2.23)	(0.41)
City x ln Population			0.0162	-0.00471
			(2.55)	(-0.50)
Constant	0.987	0.988	0.976	0.00569
	(239.43)	(221.85)	(141.38)	(0.86)
R-squared	0.366	0.244	0.281	0.081
R-MSE	0.0371	0.0403	0.0398	0.0406

Five points of reference

1. Agglomeration benefits in the higher end of distribution
2. ... which are located in cities
3. ... more likely to be related to job characteristics than to HC
4. Interregional differences in the return to HC
5. ... which are particularly high in cities

Static spatial equilibrium model

- Regions indexed r with 2 exogenous characteristics
  - mean occupational complexity O_r
  - amenity: January temperature T_r
- Workers endowed with HC h and 1 unit of labour supply
- Perfect competition/zero profit condition, no capital
- Costless interregional labour mobility
- Nationwide benchmark utility u_h for type h
- Regional commodity O_r traded on national market
- Model has 4 blocks
  1. Worker utility
  2. Regional labour markets
  3. Regional land markets
  4. Agglomeration and spatial structure

1. Worker utility

- Nation wide utility type h normalized WLOG

$$u_h = h$$

- Cost function (Talyor expansion of Comin et.al. 2015)

$$\underbrace{\omega_r + \beta_r h}_{\text{cost = income}} = \underbrace{h}_{\text{benchmark u}} + \underbrace{(1 - \varepsilon h)\,\lambda p_r}_{\text{price index}} - \underbrace{(\alpha - \xi h)'\,x_r}_{\text{public goods}}$$

  - λ: average land share in consumption
  - ε: one minus income elasticity (Albouy et.al. 2016: $\varepsilon > 0$)
  - $x_r = [T_r, \omega_r]'$: amenities (Ahlfeldt et.al. 2015)
  - p_r: log price of land

- Must hold identically for all h

$$\omega_r = \lambda p_r - \alpha' x_r$$
$$\beta_r = 1 - \varepsilon\lambda p_r + \xi' x_r$$

  - ω_r increasing land prices and decreasing in amenities
  - β_r **decreasing** in land prices (!) since land is a normal good

1. Worker utility

- Nation wide utility type h normalized WLOG

$$u_h = h$$

- Cost function (Talyor expansion of Comin et.al. 2015)

$$\underbrace{\omega_r + \beta_r h}_{\text{cost = income}} = \underbrace{h}_{\text{benchmark u}} + \underbrace{(1 - \varepsilon h)\,\lambda p_r}_{\text{price index}} - \underbrace{(\alpha - \xi h)'\, x_r}_{\text{public goods}}$$

  - λ: average land share in consumption
  - ε: one minus income elasticity (Albouy et.al. 2016: $\varepsilon > 0$)
  - $x_r = [T_r, \omega_r]'$: amenities (Ahlfeldt et.al. 2015)
  - p_r: log price of land

- Must hold identically for all h

$$\omega_r = \lambda p_r - \alpha' x_r$$

$$\beta_r = 1 - \varepsilon \lambda p_r + \xi' x_r$$

  - ω_r increasing land prices and decreasing in amenities
  - β_r **decreasing** in land prices (!) since land is a normal good

1. Worker utility

- Nation wide utility type h normalized WLOG

$$u_h = h$$

- Cost function (Talyor expansion of Comin et.al. 2015)

$$\underbrace{\omega_r + \beta_r h}_{\text{cost} = \text{income}} = \underbrace{h}_{\text{benchmark u}} + \underbrace{(1 - \varepsilon h)\,\lambda p_r}_{\text{price index}} - \underbrace{(\alpha - \xi h)'\,x_r}_{\text{public goods}}$$

  - λ: average land share in consumption
  - ε: one minus income elasticity (Albouy et.al. 2016: $\varepsilon > 0$)
  - $x_r = [T_r, \omega_r]'$: amenities (Ahlfeldt et.al. 2015)
  - p_r: log price of land

- Must hold identically for all h

$$\omega_r = \lambda p_r - \alpha' x_r$$
$$\beta_r = 1 - \varepsilon \lambda p_r + \xi' x_r$$

  - ω_r increasing land prices and decreasing in amenities
  - β_r **decreasing** in land prices (!) since land is a normal good

1. Worker utility

- Nation wide utility type h normalized WLOG

$$u_h = h$$

- Cost function (Talyor expansion of Comin et.al. 2015)

$$\underbrace{\omega_r + \beta_r h}_{\text{cost} = \text{income}} = \underbrace{h}_{\text{benchmark u}} + \underbrace{(1 - \varepsilon h)\,\lambda p_r}_{\text{price index}} - \underbrace{(\alpha - \xi h)'\, x_r}_{\text{public goods}}$$

  - λ: average land share in consumption
  - ε: one minus income elasticity (Albouy et.al. 2016: $\varepsilon > 0$)
  - $x_r = [T_r, \omega_r]'$: amenities (Ahlfeldt et.al. 2015)
  - p_r: log price of land

- Must hold identically for all h

$$\omega_r = \lambda p_r - \alpha' x_r$$

$$\beta_r = 1 - \varepsilon \lambda p_r + \xi' x_r$$

  - ω_r increasing land prices and decreasing in amenities
  - β_r **decreasing** in land prices (!) since land is a normal good

2. Regional labour markets

- Log output of worker type h in occupation o

$$y(h,o) = h - o - \frac{1}{2}\gamma(h-o)^2$$

  2 assumptions

  1. $y_h(h,o) > 0$: absolute advantage
  2. $y_{ho}(h,o) > 0$: Ricardian comparative advantage

- Output occupation o input for regional composite commodity O_r
- Cost minimization

$$h_r(o) = \arg\min_h [w_r(h) - y(h,o)]$$

- Determines equilibrium assignment and return

$$\begin{aligned} h_r(o) &= o + (H_r - O_r) \\ 1 - \beta_r &= \gamma(H_r - O_r) = \varepsilon\omega_r + (\varepsilon\alpha - \xi)' x_r. \end{aligned}$$

  - H_r and O_r are regional means of h and o
  - Optimal assignment is a mean shifter

2. Regional labour markets

- Log output of worker type h in occupation o

$$y(h,o) = h - o - \frac{1}{2}\gamma (h-o)^2$$

2 assumptions

1. $y_h(h,o) > 0$: absolute advantage
2. $y_{ho}(h,o) > 0$: Ricardian comparative advantage

- Output occupation o input for regional composite commodity O_r
- Cost minimization

$$h_r(o) = \arg\min_h \left[w_r(h) - y(h,o) \right]$$

- Determines equilibrium assignment and return

$$h_r(o) = o + (H_r - O_r)$$
$$1 - \beta_r = \gamma (H_r - O_r) = \varepsilon \omega_r + (\varepsilon \alpha - \xi)' x_r.$$

  - H_r and O_r are regional means of h and o
  - Optimal assignment is a mean shifter

2. Regional labour markets

- Log output of worker type h in occupation o

$$y\left(h,o\right)=h-o-\frac{1}{2}\gamma\left(h-o\right)^{2}$$

  2 assumptions
  1. $y_h\left(h,o\right)>0$: absolute advantage
  2. $y_{ho}\left(h,o\right)>0$: Ricardian comparative advantage

- Output occupation o input for regional composite commodity O_r
- Cost minimization

$$h_r\left(o\right)=\arg\min_h\left[w_r\left(h\right)-y\left(h,o\right)\right]$$

- Determines equilibrium assignment and return

$$\begin{aligned}h_r\left(o\right)&=o+\left(H_r-O_r\right)\\1-\beta_r&=\gamma\left(H_r-O_r\right)=\varepsilon\omega_r+\left(\varepsilon\alpha-\xi\right)'x_r.\end{aligned}$$

  - H_r and O_r are regional means of h and o
  - Optimal assignment is a mean shifter

2. Regional labour markets

- Log output of worker type h in occupation o

$$y(h,o) = h - o - \frac{1}{2}\gamma(h-o)^2$$

2 assumptions

1. $y_h(h,o) > 0$: absolute advantage
2. $y_{ho}(h,o) > 0$: Ricardian comparative advantage

- Output occupation o input for regional composite commodity O_r
- Cost minimization

$$h_r(o) = \arg\min_h [w_r(h) - y(h,o)]$$

- Determines equilibrium assignment and return

$$h_r(o) = o + (H_r - O_r)$$

$$1 - \beta_r = \gamma(H_r - O_r) = \varepsilon\omega_r + (\varepsilon\alpha - \xi)' x_r.$$

  - H_r and O_r are regional means of h and o
  - Optimal assignment is a mean shifter

2. Regional labour markets

- Log output of worker type h in occupation o

$$y(h,o) = h - o - \frac{1}{2}\gamma(h-o)^2$$

2 assumptions

1. $y_h(h,o) > 0$: absolute advantage
2. $y_{ho}(h,o) > 0$: Ricardian comparative advantage

- Output occupation o input for regional composite commodity O_r
- Cost minimization

$$h_r(o) = \arg\min_h \left[w_r(h) - y(h,o)\right]$$

- Determines equilibrium assignment and return

$$h_r(o) = o + (H_r - O_r)$$
$$1 - \beta_r = \gamma(H_r - O_r) = \varepsilon\omega_r + (\varepsilon\alpha - \xi)' x_r.$$

  - H_r and O_r are regional means of h and o
  - Optimal assignment is a mean shifter

3. Regional land markets

- Individual log land demand

$$l_r\left(h\right)=\ln\lambda+\bar{\varepsilon}h-\eta\overline{\lambda}p_r-\alpha' x_r$$

  - η: elasticity of substitution land vs. other consumption

- log average value plot of land

$$v_r\equiv l_r\left(H_r\right)+p_r=\ln\lambda+\bar{\varepsilon}H_r+\left(1-\eta\overline{\lambda}\right)p_r-\alpha' x_r$$

- For homothetic Cobb Douglas utility: $\varepsilon=0,\eta=1$

$$v_r=\underbrace{\ln\lambda}_{\text{log land share}}+\underbrace{H_r}_{\text{av.log income}}+\underbrace{\lambda p_r}_{\text{price index}}-\underbrace{\alpha' x_r}_{\text{amenities}}.$$

3. Regional land markets

- Individual log land demand

$$l_r(h) = \ln \lambda + \bar{\varepsilon} h - \eta \bar{\lambda} p_r - \alpha' x_r$$

  - η: elasticity of substitution land vs. other consumption

- log average value plot of land

$$v_r \equiv l_r(H_r) + p_r = \ln \lambda + \bar{\varepsilon} H_r + \left(1 - \eta \bar{\lambda}\right) p_r - \alpha' x_r$$

- For homothetic Cobb Douglas utility: $\varepsilon = 0, \eta = 1$

$$v_r = \underbrace{\ln \lambda}_{\text{log land share}} + \underbrace{H_r}_{\text{av.log income}} + \underbrace{\lambda p_r}_{\text{price index}} - \underbrace{\alpha' x_r}_{\text{amenities}}.$$

3. Regional land markets

- Individual log land demand

$$l_r\left(h\right) = \ln\lambda + \bar{\varepsilon}h - \eta\overline{\lambda}p_r - \alpha' x_r$$

  - η: elasticity of substitution land vs. other consumption

- log average value plot of land

$$v_r \equiv l_r\left(H_r\right) + p_r = \ln\lambda + \bar{\varepsilon}H_r + \left(1 - \eta\overline{\lambda}\right)p_r - \alpha' x_r$$

- For homothetic Cobb Douglas utility: $\varepsilon = 0, \eta = 1$

$$v_r = \underbrace{\ln\lambda}_{\text{log land share}} + \underbrace{H_r}_{\text{av.log income}} + \underbrace{\lambda p_r}_{\text{price index}} - \underbrace{\alpha' x_r}_{\text{amenities}}.$$

4. Agglomeration and spatial structure (1)

- Agglomeration

$$\omega_r = \psi \left(n_r + \theta O_r \right)$$

  - n_r: log number of workers at "a location"
  - ψ: overall agglomeration benefits
  - θ: multiplier for complex occupations (Gennaioli et al. 2013: $\theta > 1$)
- Spatial structure agglomeration
  - Lucas & Rossi-Hansberg 2002, Rossi-Hansberg & Wright 2007
  - fraction ideas surviving at distance s: $1 - \delta s$
  - fraction labour time surviving after commute s: $1 - \kappa s, \kappa < \delta$
- Rural regions
  - people stay, ideas travel
- Urban regions
  - Jobs concentrated in CBD (uses no land)
  - ideas stay, people travel

4. Agglomeration and spatial structure (1)

- Agglomeration

$$\omega_r = \psi \left(n_r + \theta O_r \right)$$

  - n_r: log number of workers at "a location"
  - ψ: overall agglomeration benefits
  - θ: multiplier for complex occupations (Gennaioli et al. 2013: $\theta > 1$)
- Spatial structure agglomeration
  - Lucas & Rossi-Hansberg 2002, Rossi-Hansberg & Wright 2007
  - fraction ideas surviving at distance s: $1 - \delta s$
  - fraction labour time surviving after commute s: $1 - \kappa s, \kappa < \delta$
- Rural regions
  - people stay, ideas travel
- Urban regions
  - Jobs concentrated in CBD (uses no land)
  - ideas stay, people travel

4. Agglomeration and spatial structure (1)

- Agglomeration

$$\omega_r = \psi\left(n_r + \theta O_r\right)$$

  - n_r: log number of workers at "a location"
  - ψ: overall agglomeration benefits
  - θ: multiplier for complex occupations (Gennaioli et al. 2013: $\theta > 1$)

- Spatial structure agglomeration
  - Lucas & Rossi-Hansberg 2002, Rossi-Hansberg & Wright 2007
  - fraction ideas surviving at distance s: $1 - \delta s$
  - fraction labour time surviving after commute s: $1 - \kappa s, \kappa < \delta$

- Rural regions
  - people stay, ideas travel

- Urban regions
  - Jobs concentrated in CBD (uses no land)
  - ideas stay, people travel

4. Agglomeration and spatial structure (1)

- Agglomeration

$$\omega_r = \psi \left(n_r + \theta O_r \right)$$

  - n_r: log number of workers at "a location"
  - ψ: overall agglomeration benefits
  - θ: multiplier for complex occupations (Gennaioli et al. 2013: $\theta > 1$)

- Spatial structure agglomeration
  - Lucas & Rossi-Hansberg 2002, Rossi-Hansberg & Wright 2007
  - fraction ideas surviving at distance s: $1 - \delta s$
  - fraction labour time surviving after commute s: $1 - \kappa s, \kappa < \delta$

- Rural regions
  - people stay, ideas travel

- Urban regions
  - Jobs concentrated in CBD (uses no land)
  - ideas stay, people travel

4. Agglomeration and spatial structure (1)

- Agglomeration

$$\omega_r = \psi \left(n_r + \theta O_r \right)$$

  - n_r: log number of workers at "a location"
  - ψ: overall agglomeration benefits
  - θ: multiplier for complex occupations (Gennaioli et al. 2013: $\theta > 1$)

- Spatial structure agglomeration
  - Lucas & Rossi-Hansberg 2002, Rossi-Hansberg & Wright 2007
  - fraction ideas surviving at distance s: $1 - \delta s$
  - fraction labour time surviving after commute s: $1 - \kappa s, \kappa < \delta$

- Rural regions
  - people stay, ideas travel

- Urban regions
  - Jobs concentrated in CBD (uses no land)
  - ideas stay, people travel

Rural areas

- Agglomeration spillover

$$\omega_r^r = \psi \ln \left(\int_0^{\delta^{-1}} 2\pi s \left(1 - \delta s\right) e^{\theta O_r - l_r} \mathrm{d}s \right)$$
$$= \psi \left(\ln \frac{\pi}{3} - 2 \ln \delta + \theta O_r - l_r \right)$$

  - l_r: lot size ($=$ inverse population density)
- Solving for ω_r^r

$$\omega_r^r = \Psi^r \left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r \right]$$

  - $\Psi^r > \psi$: reflects substitution elasticity η
- Rural spatial structure can be infinitely extended **(!)**

Rural areas

- Agglomeration spillover

$$\omega_r^r = \psi \ln \left(\int_0^{\delta^{-1}} 2\pi s \left(1 - \delta s\right) e^{\theta O_r - l_r} \mathrm{d}s \right)$$
$$= \psi \left(\ln \frac{\pi}{3} - 2 \ln \delta + \theta O_r - l_r \right)$$

  - l_r: lot size ($=$ inverse population density)

- Solving for ω_r^r

$$\omega_r^r = \Psi^r \left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r \right]$$

  - $\Psi^r > \psi$: reflects substitution elasticity η

- Rural spatial structure can be infinitely extended **(!)**

Rural areas

- Agglomeration spillover

$$\omega_r^r = \psi \ln \left(\int_0^{\delta^{-1}} 2\pi s \left(1 - \delta s\right) e^{\theta O_r - l_r} \mathrm{d}s \right)$$
$$= \psi \left(\ln \frac{\pi}{3} - 2 \ln \delta + \theta O_r - l_r \right)$$

  - l_r: lot size (= inverse population density)
- Solving for ω_r^r

$$\omega_r^r = \Psi^r \left[\psi_0 + \left(\theta - \bar{\varepsilon} \right) O_r + \psi_T T_r \right]$$

  - $\Psi^r > \psi$: reflects substitution elasticity η
- Rural spatial structure can be infinitely extended **(!)**

Rural areas

- Agglomeration spillover

$$\omega_r^r = \psi \ln \left(\int_0^{\delta^{-1}} 2\pi s \left(1 - \delta s\right) e^{\theta O_r - l_r} \mathrm{d}s \right)$$
$$= \psi \left(\ln \frac{\pi}{3} - 2 \ln \delta + \theta O_r - l_r \right)$$

  - l_r: lot size ($=$ inverse population density)

- Solving for ω_r^r

$$\omega_r^r = \Psi^r \left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r \right]$$

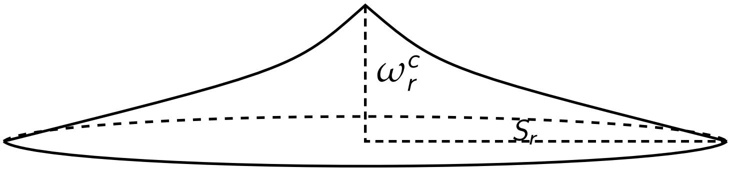
  - $\Psi^r > \psi$: reflects substitution elasticity η

- Rural spatial structure can be infinitely extended **(!)**

Urban areas

- City diameter determined by entry condition

$$\omega_r^c = -\ln(1 - \kappa S_r) \cong \kappa S_r$$



- Agglomeration spillover by a first order Taylor expansion

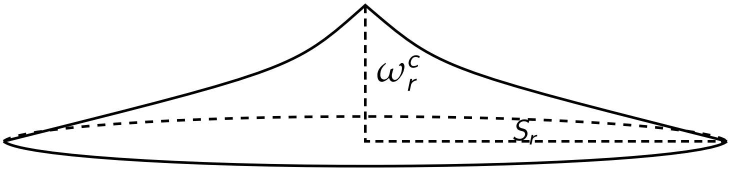
$$\omega_r^c \simeq \Psi^c \left[\psi_0 + 2\ln\left(\frac{\delta}{\kappa}\right) + (\theta - \bar{\varepsilon}) O_r + \psi_T T_r \right]$$

  - $\Psi^c > \Psi^r$: since an increase in spillovers increases n_r
  - for high O_r regions, urban structure is more efficient
- Urban structure cannot be infinitely extended **(!)**
  - Explains difference in conclusion from Hsieh & Moretti (2018)

Urban areas

- City diameter determined by entry condition

$$\omega_r^c = -\ln\left(1 - \kappa S_r\right) \cong \kappa S_r$$



- Agglomeration spillover by a first order Taylor expansion

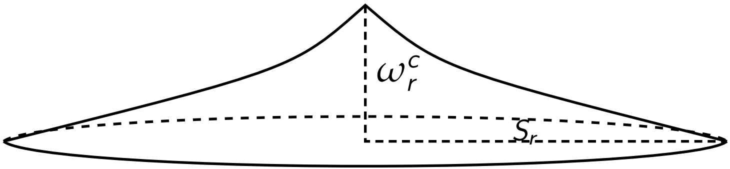
$$\omega_r^c \simeq \Psi^c \left[\psi_0 + 2\ln\left(\frac{\delta}{\kappa}\right) + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r\right]$$

  - $\Psi^c > \Psi^r$: since an increase in spillovers increases n_r
  - for high O_r regions, urban structure is more efficient
- Urban structure cannot be infinitely extended **(!)**
  - Explains difference in conclusion from Hsieh & Moretti (2018)

Urban areas

- City diameter determined by entry condition

$$\omega_r^c = -\ln\left(1 - \kappa S_r\right) \cong \kappa S_r$$



- Agglomeration spillover by a first order Taylor expansion

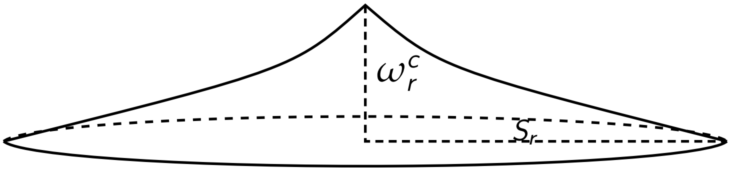
$$\omega_r^c \simeq \Psi^c \left[\psi_0 + 2\ln\left(\frac{\delta}{\kappa}\right) + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r\right]$$

  - $\Psi^c > \Psi^r$: since an increase in spillovers increases n_r
  - for high O_r regions, urban structure is more efficient
- Urban structure cannot be infinitely extended **(!)**
  - Explains difference in conclusion from Hsieh & Moretti (2018)

Urban areas

- City diameter determined by entry condition

$$\omega_r^c = -\ln(1 - \kappa S_r) \cong \kappa S_r$$



- Agglomeration spillover by a first order Taylor expansion

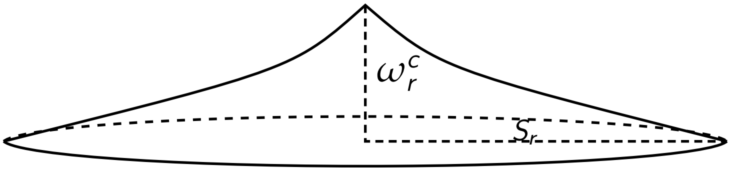
$$\omega_r^c \simeq \Psi^c \left[\psi_0 + 2\ln\left(\frac{\delta}{\kappa}\right) + (\theta - \bar{\varepsilon})\, O_r + \psi_T T_r \right]$$

  - $\Psi^c > \Psi^r$: since an increase in spillovers increases n_r
  - for high O_r regions, urban structure is more efficient
- Urban structure cannot be infinitely extended **(!)**
  - Explains difference in conclusion from Hsieh & Moretti (2018)

Urban areas

- City diameter determined by entry condition

$$\omega_r^c = -\ln\left(1 - \kappa S_r\right) \cong \kappa S_r$$



- Agglomeration spillover by a first order Taylor expansion

$$\omega_r^c \simeq \Psi^c \left[\psi_0 + 2\ln\left(\frac{\delta}{\kappa}\right) + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r\right]$$

  - $\Psi^c > \Psi^r$: since an increase in spillovers increases n_r
  - for high O_r regions, urban structure is more efficient
- Urban structure cannot be infinitely extended **(!)**
  - Explains difference in conclusion from Hsieh & Moretti (2018)

Overview of the equilibrium

$$
\begin{aligned}
1-\beta_r = \gamma\left(H_r - O_r\right) &= \varepsilon_\omega \omega_r + \varepsilon_T T_r \\
v_r &= \ln\lambda + \lambda_\omega \omega_r + \lambda_T T_r + \bar{\varepsilon} O_r \\
\omega_r^r &= \Psi^r\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r\right], \Psi^r = \frac{\psi}{1-\psi\cdot\psi_\omega} \\
\omega_r^c &= \Psi^c\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r + 2\Delta\right] \\
n_r^c &= \psi_0^+ - \bar{\varepsilon} O_r + \psi_T T_r + \psi_\omega G(\omega_r^r + 2\Psi^r\Delta) + 2\ln\left[1 - e^{-G(\omega_r^r + 2\Psi^r\Delta)}\right]
\end{aligned}
$$

Proposition

1. For $0 < 2\alpha_T < \xi_T < 4\alpha_T$ and $2\left(\alpha_\omega + 1\right) < \xi_\omega < 4\left(\alpha_\omega + 1\right)$
   - $\varepsilon_T < 0$ and $\varepsilon_\omega < 0$
   - $\lambda_T > 0$ and $\lambda_\omega > 0$
   - $\psi_T > 0$ and $\psi_\omega > 0$ (empirically, $\psi\cdot\psi_\omega < 1$)
2. For Δ sufficiently high, $\exists\omega^*$ such that $\omega^* = G\left(\omega^* + 2\Psi^r\Delta\right)$
3. $\omega_r^c > \omega_r^r$ for any $\omega_r^r > \omega^*$; $\omega_r^c < \omega_r^r$ for any $\omega_r^r < \omega^*$
4. $\Psi^c > \Psi^r$
5. $\mathrm{d}n_r^c/\mathrm{d}O_r \gtrless 0$, likely < 0 for high O_r (due to land use effect)

Overview of the equilibrium

$$
\begin{aligned}
1-\beta_r &= \gamma\left(H_r - O_r\right) = \varepsilon_\omega \omega_r + \varepsilon_T T_r \\
v_r &= \ln\lambda + \lambda_\omega \omega_r + \lambda_T T_r + \bar{\varepsilon} O_r \\
\omega_r^r &= \Psi^r\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r\right], \Psi^r = \frac{\psi}{1-\psi\cdot\psi_\omega} \\
\omega_r^c &= \Psi^c\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r + 2\Delta\right] \\
n_r^c &= \psi_0^+ - \bar{\varepsilon} O_r + \psi_T T_r + \psi_\omega G(\omega_r^r + 2\Psi^r\Delta) + 2\ln\left[1 - e^{-G(\omega_r^r + 2\Psi^r\Delta)}\right]
\end{aligned}
$$

Proposition

1. For $0 < 2\alpha_T < \xi_T < 4\alpha_T$ and $2\left(\alpha_\omega + 1\right) < \xi_\omega < 4\left(\alpha_\omega + 1\right)$
   - $\varepsilon_T < 0$ and $\varepsilon_\omega < 0$
   - $\lambda_T > 0$ and $\lambda_\omega > 0$
   - $\psi_T > 0$ and $\psi_\omega > 0$ (empirically, $\psi \cdot \psi_\omega < 1$)
2. For Δ sufficiently high, $\exists\omega^*$ such that $\omega^* = G\left(\omega^* + 2\Psi^r\Delta\right)$
3. $\omega_r^c > \omega_r^r$ for any $\omega_r^r > \omega^*$; $\omega_r^c < \omega_r^r$ for any $\omega_r^r < \omega^*$
4. $\Psi^c > \Psi^r$
5. $\mathrm{d}n_r^c/\mathrm{d}O_r \gtrless 0$, likely < 0 for high O_r (due to land use effect)

Overview of the equilibrium

$$
\begin{aligned}
1-\beta_r = \gamma\left(H_r - O_r\right) &= \varepsilon_\omega \omega_r + \varepsilon_T T_r \\
v_r &= \ln\lambda + \lambda_\omega \omega_r + \lambda_T T_r + \bar{\varepsilon} O_r \\
\omega_r^r &= \Psi^r\left[\psi_0 + \left(\theta-\bar{\varepsilon}\right) O_r + \psi_T T_r\right], \Psi^r = \frac{\psi}{1-\psi\cdot\psi_\omega} \\
\omega_r^c &= \Psi^c\left[\psi_0 + \left(\theta-\bar{\varepsilon}\right) O_r + \psi_T T_r + 2\Delta\right] \\
n_r^c &= \psi_0^+ - \bar{\varepsilon} O_r + \psi_T T_r + \psi_\omega G(\omega_r^r + 2\Psi^r\Delta) + 2\ln\left[1 - e^{-G(\omega_r^r + 2\Psi^r\Delta)}\right]
\end{aligned}
$$

Proposition

1. For $0 < 2\alpha_T < \xi_T < 4\alpha_T$ and $2\left(\alpha_\omega + 1\right) < \xi_\omega < 4\left(\alpha_\omega + 1\right)$
   - $\varepsilon_T < 0$ and $\varepsilon_\omega < 0$
   - $\lambda_T > 0$ and $\lambda_\omega > 0$
   - $\psi_T > 0$ and $\psi_\omega > 0$ (empirically, $\psi\cdot\psi_\omega < 1$)
2. For Δ sufficiently high, $\exists\omega^*$ such that $\omega^* = G\left(\omega^* + 2\Psi^r\Delta\right)$
3. $\omega_r^c > \omega_r^r$ for any $\omega_r^r > \omega^*$; $\omega_r^c < \omega_r^r$ for any $\omega_r^r < \omega^*$
4. $\Psi^c > \Psi^r$
5. $\mathrm{d}n_r^c/\mathrm{d}O_r \gtrless 0$, likely < 0 for high O_r (due to land use effect)

Overview of the equilibrium

$$
\begin{aligned}
1-\beta_r &= \gamma\left(H_r - O_r\right) = \varepsilon_\omega \omega_r + \varepsilon_T T_r \\
v_r &= \ln\lambda + \lambda_\omega \omega_r + \lambda_T T_r + \bar{\varepsilon} O_r \\
\omega_r^r &= \Psi^r\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r\right], \Psi^r = \frac{\psi}{1-\psi\cdot\psi_\omega} \\
\omega_r^c &= \Psi^c\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r + 2\Delta\right] \\
n_r^c &= \psi_0^+ - \bar{\varepsilon} O_r + \psi_T T_r + \psi_\omega G(\omega_r^r + 2\Psi^r\Delta) + 2\ln\left[1 - e^{-G(\omega_r^r + 2\Psi^r\Delta)}\right]
\end{aligned}
$$

Proposition

1. For $0 < 2\alpha_T < \xi_T < 4\alpha_T$ and $2\left(\alpha_\omega + 1\right) < \xi_\omega < 4\left(\alpha_\omega + 1\right)$
    - $\varepsilon_T < 0$ and $\varepsilon_\omega < 0$
    - $\lambda_T > 0$ and $\lambda_\omega > 0$
    - $\psi_T > 0$ and $\psi_\omega > 0$ (empirically, $\psi \cdot \psi_\omega < 1$)
2. For Δ sufficiently high, $\exists\omega^*$ such that $\omega^* = G\left(\omega^* + 2\Psi^r\Delta\right)$
3. $\omega_r^c > \omega_r^r$ for any $\omega_r^r > \omega^*$; $\omega_r^c < \omega_r^r$ for any $\omega_r^r < \omega^*$
4. $\Psi^c > \Psi^r$
5. $\mathrm{d}n_r^c/\mathrm{d}O_r \gtrless 0$, likely < 0 for high O_r (due to land use effect)

Overview of the equilibrium

$$
\begin{aligned}
1-\beta_r = \gamma\left(H_r - O_r\right) &= \varepsilon_\omega \omega_r + \varepsilon_T T_r \\
v_r &= \ln\lambda + \lambda_\omega \omega_r + \lambda_T T_r + \bar{\varepsilon} O_r \\
\omega_r^r &= \Psi^r\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r\right], \Psi^r = \frac{\psi}{1-\psi\cdot\psi_\omega} \\
\omega_r^c &= \Psi^c\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r + 2\Delta\right] \\
n_r^c &= \psi_0^+ - \bar{\varepsilon} O_r + \psi_T T_r + \psi_\omega G\left(\omega_r^r + 2\Psi^r\Delta\right) + 2\ln\left[1 - e^{-G\left(\omega_r^r + 2\Psi^r\Delta\right)}\right]
\end{aligned}
$$

Proposition

1. For $0 < 2\alpha_T < \xi_T < 4\alpha_T$ and $2\left(\alpha_\omega + 1\right) < \xi_\omega < 4\left(\alpha_\omega + 1\right)$
   - $\varepsilon_T < 0$ and $\varepsilon_\omega < 0$
   - $\lambda_T > 0$ and $\lambda_\omega > 0$
   - $\psi_T > 0$ and $\psi_\omega > 0$ (empirically, $\psi\cdot\psi_\omega < 1$)
2. For Δ sufficiently high, $\exists\omega^*$ such that $\omega^* = G\left(\omega^* + 2\Psi^r\Delta\right)$
3. $\omega_r^c > \omega_r^r$ for any $\omega_r^r > \omega^*$; $\omega_r^c < \omega_r^r$ for any $\omega_r^r < \omega^*$
4. $\Psi^c > \Psi^r$
5. $\mathrm{d}n_r^c/\mathrm{d}O_r \gtrless 0$, likely < 0 for high O_r (due to land use effect)

Overview of the equilibrium

$$1 - \beta_r = \gamma\left(H_r - O_r\right) = \varepsilon_\omega \omega_r + \varepsilon_T T_r$$

$$v_r = \ln \lambda + \lambda_\omega \omega_r + \lambda_T T_r + \bar{\varepsilon} O_r$$

$$\omega_r^r = \Psi^r\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r\right], \Psi^r = \frac{\psi}{1 - \psi \cdot \psi_\omega}$$

$$\omega_r^c = \Psi^c\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r + 2\Delta\right]$$

$$n_r^c = \psi_0^+ - \bar{\varepsilon} O_r + \psi_T T_r + \psi_\omega G(\omega_r^r + 2\Psi^r \Delta) + 2\ln\left[1 - e^{-G(\omega_r^r + 2\Psi^r\Delta)}\right]$$

Proposition

1. For $0 < 2\alpha_T < \xi_T < 4\alpha_T$ and $2\left(\alpha_\omega + 1\right) < \xi_\omega < 4\left(\alpha_\omega + 1\right)$
   - $\varepsilon_T < 0$ and $\varepsilon_\omega < 0$
   - $\lambda_T > 0$ and $\lambda_\omega > 0$
   - $\psi_T > 0$ and $\psi_\omega > 0$ (empirically, $\psi \cdot \psi_\omega < 1$)
2. For Δ sufficiently high, $\exists \omega^*$ such that $\omega^* = G\left(\omega^* + 2\Psi^r\Delta\right)$
3. $\omega_r^c > \omega_r^r$ for any $\omega_r^r > \omega^*$; $\omega_r^c < \omega_r^r$ for any $\omega_r^r < \omega^*$
4. $\Psi^c > \Psi^r$
5. $\mathrm{d}n_r^c / \mathrm{d}O_r \gtrless 0$, likely < 0 for high O_r (due to land use effect)

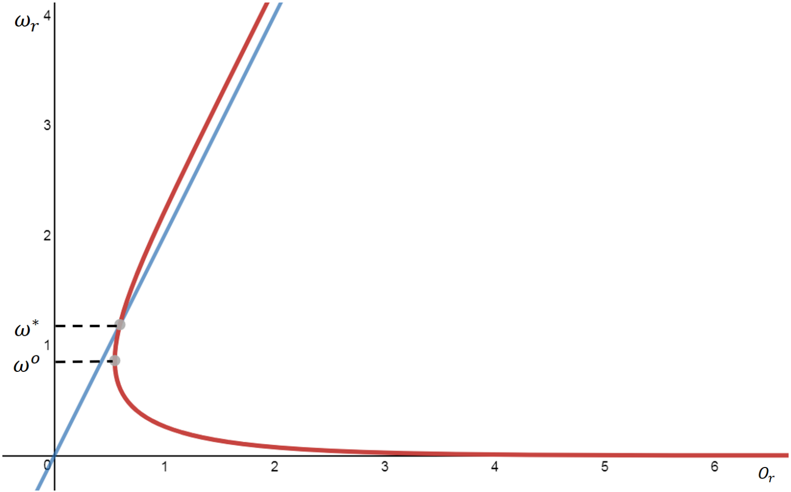
Overview of the equilibrium

$$1-\beta_r = \gamma\left(H_r - O_r\right) = \varepsilon_\omega \omega_r + \varepsilon_T T_r$$
$$v_r = \ln\lambda + \lambda_\omega \omega_r + \lambda_T T_r + \bar{\varepsilon} O_r$$
$$\omega_r^r = \Psi^r\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r\right], \Psi^r = \frac{\psi}{1-\psi\cdot\psi_\omega}$$
$$\omega_r^c = \Psi^c\left[\psi_0 + \left(\theta - \bar{\varepsilon}\right) O_r + \psi_T T_r + 2\Delta\right]$$
$$n_r^c = \psi_0^+ - \bar{\varepsilon} O_r + \psi_T T_r + \psi_\omega G(\omega_r^r + 2\Psi^r\Delta) + 2\ln\left[1 - e^{-G(\omega_r^r + 2\Psi^r\Delta)}\right]$$

Proposition

1. For $0 < 2\alpha_T < \xi_T < 4\alpha_T$ and $2\left(\alpha_\omega + 1\right) < \xi_\omega < 4\left(\alpha_\omega + 1\right)$
    - $\varepsilon_T < 0$ and $\varepsilon_\omega < 0$
    - $\lambda_T > 0$ and $\lambda_\omega > 0$
    - $\psi_T > 0$ and $\psi_\omega > 0$ (empirically, $\psi\cdot\psi_\omega < 1$)
2. For Δ sufficiently high, $\exists\omega^*$ such that $\omega^* = G\left(\omega^* + 2\Psi^r\Delta\right)$
3. $\omega_r^c > \omega_r^r$ for any $\omega_r^r > \omega^*$; $\omega_r^c < \omega_r^r$ for any $\omega_r^r < \omega^*$
4. $\Psi^c > \Psi^r$
5. $\mathrm{d}n_r^c/\mathrm{d}O_r \gtrless 0$, likely < 0 for high O_r (due to land use effect)

Illustration



Agglomeration of City ω_r^c and non-city ω_r^r

5. Commodity market and spatial structure

- MISSING in this paper
- Rossi-Hansberg & Wright (2007)
- Free entry of regions in each O_r market
- Drives commodity prices down to production cost
- ... potentially with remaining inefficiencies due to knowledge spillovers
  - that is not in Rossi-Hansberg & Wright (2007)
- Drives spatial structure O_r or the other way round?
  - Glaeser (2005): reinventing the city
- Market for spatial structure
  - irreversibility of urbanization

Empirical implementation

- Instrument for O_r
  - Static Bartik instrument
  - Nation wide O by industry
  - Weighting that by regional industry mix

Bartik IV regression (1)

VARIABLES	(1) O_r	(2) $1-\beta_r$	(3) $H_r - O_r$	(4) v_r	(5) ω_r
Panel A: City and Non-City Sample (81 Observations)					
ln Jan Temp	-0.00585	-0.0270	-0.0566	0.161	0.0438
	(-1.03)	(-2.36)	(-5.50)	(1.78)	(2.99)
Bartik IV	1.524				
	(19.41)				
Bartik IV sq.	3.698				
	(3.40)				
Occ IV		-0.384	-0.282	3.110	0.536
		(-3.77)	(-3.06)	(3.86)	(4.10)
Metro Dummy	0.0246	-0.00567	0.0146	-0.0943	0.0262
	(4.63)	(-0.46)	(1.33)	(-0.98)	(1.67)
Constant	0.0217	0.171	0.321	10.70	-0.330
	(0.67)	(2.64)	(5.50)	(20.96)	(-3.97)
R-squared	0.911	0.367	0.368	0.251	0.504
R-MSE	0.0185	0.0373	0.0337	0.295	0.0479

Bartik IV regression (1)

VARIABLES	(1) O_r	(2) $1-\beta_r$	(3) $H_r - O_r$	(4) v_r	(5) ω_r
Panel A: City and Non-City Sample (81 Observations)					
ln Jan Temp	-0.00585	-0.0270	-0.0566	0.161	0.0438
	(-1.03)	(-2.36)	(-5.50)	(1.78)	(2.99)
Bartik IV	1.524				
	(19.41)				
Bartik IV sq.	3.698				
	(3.40)				
Occ IV		-0.384	-0.282	3.110	0.536
		(-3.77)	(-3.06)	(3.86)	(4.10)
Metro Dummy	0.0246	-0.00567	0.0146	-0.0943	0.0262
	(4.63)	(-0.46)	(1.33)	(-0.98)	(1.67)
Constant	0.0217	0.171	0.321	10.70	-0.330
	(0.67)	(2.64)	(5.50)	(20.96)	(-3.97)
R-squared	0.911	0.367	0.368	0.251	0.504
R-MSE	0.0185	0.0373	0.0337	0.295	0.0479

Bartik IV regression (1)

VARIABLES	(1) O_r	(2) $1-\beta_r$	(3) $H_r - O_r$	(4) v_r	(5) ω_r
Panel A: City and Non-City Sample (81 Observations)					
ln Jan Temp	-0.00585	-0.0270	-0.0566	0.161	0.0438
	(-1.03)	(-2.36)	(-5.50)	(1.78)	(2.99)
Bartik IV	1.524				
	(19.41)				
Bartik IV sq.	3.698				
	(3.40)				
Occ IV		-0.384	-0.282	3.110	0.536
		(-3.77)	(-3.06)	(3.86)	(4.10)
Metro Dummy	0.0246	-0.00567	0.0146	-0.0943	0.0262
	(4.63)	(-0.46)	(1.33)	(-0.98)	(1.67)
Constant	0.0217	0.171	0.321	10.70	-0.330
	(0.67)	(2.64)	(5.50)	(20.96)	(-3.97)
R-squared	0.911	0.367	0.368	0.251	0.504
R-MSE	0.0185	0.0373	0.0337	0.295	0.0479

Bartik IV regression (1)

VARIABLES	(1) O_r	(2) $1-\beta_r$	(3) $H_r - O_r$	(4) v_r	(5) ω_r
Panel A: City and Non-City Sample (81 Observations)					
ln Jan Temp	-0.00585	-0.0270	-0.0566	0.161	0.0438
	(-1.03)	(-2.36)	(-5.50)	(1.78)	(2.99)
Bartik IV	1.524				
	(19.41)				
Bartik IV sq.	3.698				
	(3.40)				
Occ IV		-0.384	-0.282	3.110	0.536
		(-3.77)	(-3.06)	(3.86)	(4.10)
Metro Dummy	0.0246	-0.00567	0.0146	-0.0943	0.0262
	(4.63)	(-0.46)	(1.33)	(-0.98)	(1.67)
Constant	0.0217	0.171	0.321	10.70	-0.330
	(0.67)	(2.64)	(5.50)	(20.96)	(-3.97)
R-squared	0.911	0.367	0.368	0.251	0.504
R-MSE	0.0185	0.0373	0.0337	0.295	0.0479

Bartik IV regression (1)

VARIABLES	(1) O_r	(2) $1-\beta_r$	(3) $H_r - O_r$	(4) v_r	(5) ω_r
Panel A: City and Non-City Sample (81 Observations)					
ln Jan Temp	-0.00585	-0.0270	-0.0566	0.161	0.0438
	(-1.03)	(-2.36)	(-5.50)	(1.78)	(2.99)
Bartik IV	1.524				
	(19.41)				
Bartik IV sq.	3.698				
	(3.40)				
Occ IV		-0.384	-0.282	3.110	0.536
		(-3.77)	(-3.06)	(3.86)	(4.10)
Metro Dummy	0.0246	-0.00567	0.0146	-0.0943	0.0262
	(4.63)	(-0.46)	(1.33)	(-0.98)	(1.67)
Constant	0.0217	0.171	0.321	10.70	-0.330
	(0.67)	(2.64)	(5.50)	(20.96)	(-3.97)
R-squared	0.911	0.367	0.368	0.251	0.504
R-MSE	0.0185	0.0373	0.0337	0.295	0.0479

Bartik IV regression (1)

VARIABLES	(1) O_r	(2) $1-\beta_r$	(3) $H_r - O_r$	(4) v_r	(5) ω_r
Panel A: City and Non-City Sample (81 Observations)					
ln Jan Temp	-0.00585	-0.0270	-0.0566	0.161	0.0438
	(-1.03)	(-2.36)	(-5.50)	(1.78)	(2.99)
Bartik IV	1.524				
	(19.41)				
Bartik IV sq.	3.698				
	(3.40)				
Occ IV		-0.384	-0.282	3.110	0.536
		(-3.77)	(-3.06)	(3.86)	(4.10)
Metro Dummy	0.0246	-0.00567	0.0146	-0.0943	0.0262
	(4.63)	(-0.46)	(1.33)	(-0.98)	(1.67)
Constant	0.0217	0.171	0.321	10.70	-0.330
	(0.67)	(2.64)	(5.50)	(20.96)	(-3.97)
R-squared	0.911	0.367	0.368	0.251	0.504
R-MSE	0.0185	0.0373	0.0337	0.295	0.0479

Bartik IV regression (1)

VARIABLES	(1) O_r	(2) $1-\beta_r$	(3) $H_r - O_r$	(4) v_r	(5) ω_r
Panel A: City and Non-City Sample (81 Observations)					
ln Jan Temp	-0.00585	-0.0270	-0.0566	0.161	0.0438
	(-1.03)	(-2.36)	(-5.50)	(1.78)	(2.99)
Bartik IV	1.524				
	(19.41)				
Bartik IV sq.	3.698				
	(3.40)				
Occ IV		-0.384	-0.282	3.110	0.536
		(-3.77)	(-3.06)	(3.86)	(4.10)
Metro Dummy	0.0246	-0.00567	0.0146	-0.0943	0.0262
	(4.63)	(-0.46)	(1.33)	(-0.98)	(1.67)
Constant	0.0217	0.171	0.321	10.70	-0.330
	(0.67)	(2.64)	(5.50)	(20.96)	(-3.97)
R-squared	0.911	0.367	0.368	0.251	0.504
R-MSE	0.0185	0.0373	0.0337	0.295	0.0479

Bartik IV regression (2)

VARIABLES	(1) v_r	(2) ω_r
Panel B: City Sample (34 Observations)		
ln Jan Temp	0.476	0.0387
	(3.12)	(1.79)
Occ IV	5.016	0.675
	(4.54)	(4.29)
Constant	8.659	-0.280
	(9.55)	(-2.17)
R-squared	0.463	0.390
R-MSE	0.300	0.0426
Panel C: Non-city Sample (47 Observations)		
ln Jan Temp	-0.00426	0.0481
	(-0.04)	(2.44)
Occ IV	1.107	0.356
	(1.05)	(1.69)
Constant	11.57	-0.360
	(20.63)	(-3.21)
R-squared	0.025	0.164
R-MSE	0.257	0.0515

Bartik IV regression (2)

	(1)	(2)
VARIABLES	v_r	ω_r
Panel B: City Sample (34 Observations)		
ln Jan Temp	0.476	0.0387
	(3.12)	(1.79)
Occ IV	5.016	0.675
	(4.54)	(4.29)
Constant	8.659	-0.280
	(9.55)	(-2.17)
R-squared	0.463	0.390
R-MSE	0.300	0.0426
Panel C: Non-city Sample (47 Observations)		
ln Jan Temp	-0.00426	0.0481
	(-0.04)	(2.44)
Occ IV	1.107	0.356
	(1.05)	(1.69)
Constant	11.57	-0.360
	(20.63)	(-3.21)
R-squared	0.025	0.164
R-MSE	0.257	0.0515

VARIABLES	(1) $1-\beta_r$	(2) $H_r - O_r$	(3) v_r	(4) ω_r	(5) n_r
Table 7 Panel B: City Sample (34 Observations)					
ln Jan Temp	0.000898	-0.0320	0.476	0.0387	0.397
	(0.05)	(-1.85)	(3.12)	(1.79)	(1.19)
Occ IV	-0.207	-0.172	5.016	0.675	2.311
	(-1.71)	(-1.37)	(4.54)	(4.29)	(0.96)
Constant	-0.00725	0.185	8.659	-0.280	11.87
	(-0.07)	(1.80)	(9.55)	(-2.17)	(5.99)
R-squared	0.089	0.130	0.463	0.390	0.062
R-MSE	0.0327	0.0340	0.300	0.0426	0.655
Table A5 Panel B: City Sample (222 Observations)					
ln Jan Temp	-0.0165	-0.0680	0.142	0.0424	0.649
	(-1.07)	(-6.81)	(1.21)	(3.41)	(3.59)
Occ IV	-0.255	-0.144	0.864	-0.0207	3.371
	(-3.49)	(-3.04)	(1.56)	(-0.35)	(3.93)
Constant	0.107	0.390	11.08	-0.224	9.537
	(1.20)	(6.72)	(16.29)	(-3.09)	(9.06)
R-squared	0.054	0.189	0.015	0.053	0.102
R-MSE	0.0833	0.0538	0.631	0.0672	0.976

Bartik IV regression (3)

VARIABLES	(1) $1-\beta_r$	(2) $H_r - O_r$	(3) v_r	(4) ω_r	(5) n_r
Table 7 Panel B: City Sample (34 Observations)					
ln Jan Temp	0.000898	-0.0320	0.476	0.0387	0.397
	(0.05)	(-1.85)	(3.12)	(1.79)	(1.19)
Occ IV	-0.207	-0.172	5.016	0.675	2.311
	(-1.71)	(-1.37)	(4.54)	(4.29)	(0.96)
Constant	-0.00725	0.185	8.659	-0.280	11.87
	(-0.07)	(1.80)	(9.55)	(-2.17)	(5.99)
R-squared	0.089	0.130	0.463	0.390	0.062
R-MSE	0.0327	0.0340	0.300	0.0426	0.655
Table A5 Panel B: City Sample (222 Observations)					
ln Jan Temp	-0.0165	-0.0680	0.142	0.0424	0.649
	(-1.07)	(-6.81)	(1.21)	(3.41)	(3.59)
Occ IV	-0.255	-0.144	0.864	-0.0207	3.371
	(-3.49)	(-3.04)	(1.56)	(-0.35)	(3.93)
Constant	0.107	0.390	11.08	-0.224	9.537
	(1.20)	(6.72)	(16.29)	(-3.09)	(9.06)
R-squared	0.054	0.189	0.015	0.053	0.102
R-MSE	0.0833	0.0538	0.631	0.0672	0.976

Bartik IV regression (3)

VARIABLES	(1) $1-\beta_r$	(2) $H_r - O_r$	(3) v_r	(4) ω_r	(5) n_r
Table 7 Panel B: City Sample (34 Observations)					
ln Jan Temp	0.000898	-0.0320	0.476	0.0387	0.397
	(0.05)	(-1.85)	(3.12)	(1.79)	(1.19)
Occ IV	-0.207	-0.172	5.016	0.675	2.311
	(-1.71)	(-1.37)	(4.54)	(4.29)	(0.96)
Constant	-0.00725	0.185	8.659	-0.280	11.87
	(-0.07)	(1.80)	(9.55)	(-2.17)	(5.99)
R-squared	0.089	0.130	0.463	0.390	0.062
R-MSE	0.0327	0.0340	0.300	0.0426	0.655
Table A5 Panel B: City Sample (222 Observations)					
ln Jan Temp	-0.0165	-0.0680	0.142	0.0424	0.649
	(-1.07)	(-6.81)	(1.21)	(3.41)	(3.59)
Occ IV	-0.255	-0.144	0.864	-0.0207	3.371
	(-3.49)	(-3.04)	(1.56)	(-0.35)	(3.93)
Constant	0.107	0.390	11.08	-0.224	9.537
	(1.20)	(6.72)	(16.29)	(-3.09)	(9.06)
R-squared	0.054	0.189	0.015	0.053	0.102
R-MSE	0.0833	0.0538	0.631	0.0672	0.976

Ranking occupational complexity

Region	Occ Index	Type	Region	Occ Index	Type	Region	Occ Index	Type
Washington, DC	0.093	City	Virginia Beach, VA	0.013	City	Nebraska	-0.017	Non-City
San Jose, CA	0.093	City	Cincinnati, OH	0.012	City	South Carolina	-0.017	Non-City
Boston, MA	0.070	City	Massachusetts	0.011	Non-City	Michigan	-0.018	Non-City
San Francisco, CA	0.056	City	New Orleans, LA	0.011	City	Illinois	-0.018	Non-City
Seattle, WA	0.056	City	Utah	0.008	Non-City	Riverside, CA	-0.019	City
Denver, CO	0.053	City	Colorado	0.008	Non-City	Maryland	-0.019	Non-City
Baltimore, MD	0.050	City	Rhode Island	0.007	Non-City	Iowa	-0.020	Non-City
Connecticut	0.047	Non-City	Tampa, FL	0.006	City	Miami, FL	-0.020	City
Minneapolis, MN	0.041	City	Vermont	0.004	Non-City	Tennessee	-0.021	Non-City
Atlanta, GA	0.039	City	Arizona	0.002	Non-City	Kentucky	-0.023	Non-City
Philadelphia, PA	0.037	City	New Mexico	0.001	Non-City	Pennsylvania	-0.025	Non-City
Kansas City, MO	0.032	City	New York	0.000	Non-City	North Dakota	-0.025	Non-City
Indianapolis, IN	0.029	City	Oklahoma	-0.003	Non-City	Wisconsin	-0.027	Non-City
New Hampshire	0.029	Non-City	Delaware	-0.003	Non-City	Mississippi	-0.029	Non-City
Dallas, TX	0.029	City	Virginia	-0.004	Non-City	Washington	-0.029	Non-City
Chicago, IL	0.027	City	Greensboro, NC	-0.007	City	Texas	-0.030	Non-City
Portland, OR	0.026	City	Los Angeles, CA	-0.008	City	Montana	-0.031	Non-City
Houston, TX	0.024	City	Wyoming	-0.009	Non-City	California	-0.031	Non-City
Milwaukee, WI	0.024	City	Kansas	-0.012	Non-City	Indiana	-0.031	Non-City
Pittsburgh, PA	0.020	City	North Carolina	-0.012	Non-City	Idaho	-0.033	Non-City
Detroit, MI	0.020	City	Florida	-0.013	Non-City	South Dakota	-0.037	Non-City
Rochester, NY	0.019	City	Alabama	-0.014	Non-City	Georgia	-0.037	Non-City
Cleveland, OH	0.019	City	Louisiana	-0.014	Non-City	Arkansas	-0.038	Non-City
New York, NY	0.018	City	West Virginia	-0.015	Non-City	Missouri	-0.040	Non-City
St Louis, MO	0.017	City	Maine	-0.016	Non-City	Oregon	-0.044	Non-City
Columbus, OH	0.017	City	Buffalo, NY	-0.016	City	Minnesota	-0.049	Non-City
San Diego, CA	0.016	City	Ohio	-0.017	Non-City	Nevada	-0.074	Non-City

Ranking occupational complexity

Region	Occ Index	Type	Region	Occ Index	Type	Region	Occ Index	Type
Washington, DC	0.093	City	Virginia Beach, VA	0.013	City	Nebraska	-0.017	Non-City
San Jose, CA	0.093	City	Cincinnati, OH	0.012	City	South Carolina	-0.017	Non-City
Boston, MA	0.070	City	Massachusetts	0.011	Non-City	Michigan	-0.018	Non-City
San Francisco, CA	0.056	City	New Orleans, LA	0.011	City	Illinois	-0.018	Non-City
Seattle, WA	0.056	City	Utah	0.008	Non-City	Riverside, CA	-0.019	City
Denver, CO	0.053	City	Colorado	0.008	Non-City	Maryland	-0.019	Non-City
Baltimore, MD	0.050	City	Rhode Island	0.007	Non-City	Iowa	-0.020	Non-City
Connecticut	0.047	Non-City	Tampa, FL	0.006	City	Miami, FL	-0.020	City
Minneapolis, MN	0.041	City	Vermont	0.004	Non-City	Tennessee	-0.021	Non-City
Atlanta, GA	0.039	City	Arizona	0.002	Non-City	Kentucky	-0.023	Non-City
Philadelphia, PA	0.037	City	New Mexico	0.001	Non-City	Pennsylvania	-0.025	Non-City
Kansas City, MO	0.032	City	New York	0.000	Non-City	North Dakota	-0.025	Non-City
Indianapolis, IN	0.029	City	Oklahoma	-0.003	Non-City	Wisconsin	-0.027	Non-City
New Hampshire	0.029	Non-City	Delaware	-0.003	Non-City	Mississippi	-0.029	Non-City
Dallas, TX	0.029	City	Virginia	-0.004	Non-City	Washington	-0.029	Non-City
Chicago, IL	0.027	City	Greensboro, NC	-0.007	City	Texas	-0.030	Non-City
Portland, OR	0.026	City	Los Angeles, CA	-0.008	City	Montana	-0.031	Non-City
Houston, TX	0.024	City	Wyoming	-0.009	Non-City	California	-0.031	Non-City
Milwaukee, WI	0.024	City	Kansas	-0.012	Non-City	Indiana	-0.031	Non-City
Pittsburgh, PA	0.020	City	North Carolina	-0.012	Non-City	Idaho	-0.033	Non-City
Detroit, MI	0.020	City	Florida	-0.013	Non-City	South Dakota	-0.037	Non-City
Rochester, NY	0.019	City	Alabama	-0.014	Non-City	Georgia	-0.037	Non-City
Cleveland, OH	0.019	City	Louisiana	-0.014	Non-City	Arkansas	-0.038	Non-City
New York, NY	0.018	City	West Virginia	-0.015	Non-City	Missouri	-0.040	Non-City
St Louis, MO	0.017	City	Maine	-0.016	Non-City	Oregon	-0.044	Non-City
Columbus, OH	0.017	City	Buffalo, NY	-0.016	City	Minnesota	-0.049	Non-City
San Diego, CA	0.016	City	Ohio	-0.017	Non-City	Nevada	-0.074	Non-City

Ranking occupational complexity

Region	Occ Index	Type	Region	Occ Index	Type	Region	Occ Index	Type
Washington, DC	0.093	City	Virginia Beach, VA	0.013	City	Nebraska	-0.017	Non-City
San Jose, CA	0.093	City	Cincinnati, OH	0.012	City	South Carolina	-0.017	Non-City
Boston, MA	0.070	City	Massachusetts	0.011	Non-City	Michigan	-0.018	Non-City
San Francisco, CA	0.056	City	New Orleans, LA	0.011	City	Illinois	-0.018	Non-City
Seattle, WA	0.056	City	Utah	0.008	Non-City	Riverside, CA	-0.019	City
Denver, CO	0.053	City	Colorado	0.008	Non-City	Maryland	-0.019	Non-City
Baltimore, MD	0.050	City	Rhode Island	0.007	Non-City	Iowa	-0.020	Non-City
Connecticut	0.047	Non-City	Tampa, FL	0.006	City	Miami, FL	-0.020	City
Minneapolis, MN	0.041	City	Vermont	0.004	Non-City	Tennessee	-0.021	Non-City
Atlanta, GA	0.039	City	Arizona	0.002	Non-City	Kentucky	-0.023	Non-City
Philadelphia, PA	0.037	City	New Mexico	0.001	Non-City	Pennsylvania	-0.025	Non-City
Kansas City, MO	0.032	City	New York	0.000	Non-City	North Dakota	-0.025	Non-City
Indianapolis, IN	0.029	City	Oklahoma	-0.003	Non-City	Wisconsin	-0.027	Non-City
New Hampshire	0.029	Non-City	Delaware	-0.003	Non-City	Mississippi	-0.029	Non-City
Dallas, TX	0.029	City	Virginia	-0.004	Non-City	Washington	-0.029	Non-City
Chicago, IL	0.027	City	Greensboro, NC	-0.007	City	Texas	-0.030	Non-City
Portland, OR	0.026	City	Los Angeles, CA	-0.008	City	Montana	-0.031	Non-City
Houston, TX	0.024	City	Wyoming	-0.009	Non-City	California	-0.031	Non-City
Milwaukee, WI	0.024	City	Kansas	-0.012	Non-City	Indiana	-0.031	Non-City
Pittsburgh, PA	0.020	City	North Carolina	-0.012	Non-City	Idaho	-0.033	Non-City
Detroit, MI	0.020	City	Florida	-0.013	Non-City	South Dakota	-0.037	Non-City
Rochester, NY	0.019	City	Alabama	-0.014	Non-City	Georgia	-0.037	Non-City
Cleveland, OH	0.019	City	Louisiana	-0.014	Non-City	Arkansas	-0.038	Non-City
New York, NY	0.018	City	West Virginia	-0.015	Non-City	Missouri	-0.040	Non-City
St Louis, MO	0.017	City	Maine	-0.016	Non-City	Oregon	-0.044	Non-City
Columbus, OH	0.017	City	Buffalo, NY	-0.016	City	Minnesota	-0.049	Non-City
San Diego, CA	0.016	City	Ohio	-0.017	Non-City	Nevada	-0.074	Non-City

Ranking occupational complexity

Region	Occ Index	Type	Region	Occ Index	Type	Region	Occ Index	Type
Washington, DC	0.093	City	Virginia Beach, VA	0.013	City	Nebraska	-0.017	Non-City
San Jose, CA	0.093	City	Cincinnati, OH	0.012	City	South Carolina	-0.017	Non-City
Boston, MA	0.070	City	Massachusetts	0.011	Non-City	Michigan	-0.018	Non-City
San Francisco, CA	0.056	City	New Orleans, LA	0.011	City	Illinois	-0.018	Non-City
Seattle, WA	0.056	City	Utah	0.008	Non-City	Riverside, CA	-0.019	City
Denver, CO	0.053	City	Colorado	0.008	Non-City	Maryland	-0.019	Non-City
Baltimore, MD	0.050	City	Rhode Island	0.007	Non-City	Iowa	-0.020	Non-City
Connecticut	0.047	Non-City	Tampa, FL	0.006	City	Miami, FL	-0.020	City
Minneapolis, MN	0.041	City	Vermont	0.004	Non-City	Tennessee	-0.021	Non-City
Atlanta, GA	0.039	City	Arizona	0.002	Non-City	Kentucky	-0.023	Non-City
Philadelphia, PA	0.037	City	New Mexico	0.001	Non-City	Pennsylvania	-0.025	Non-City
Kansas City, MO	0.032	City	New York	0.000	Non-City	North Dakota	-0.025	Non-City
Indianapolis, IN	0.029	City	Oklahoma	-0.003	Non-City	Wisconsin	-0.027	Non-City
New Hampshire	0.029	Non-City	Delaware	-0.003	Non-City	Mississippi	-0.029	Non-City
Dallas, TX	0.029	City	Virginia	-0.004	Non-City	Washington	-0.029	Non-City
Chicago, IL	0.027	City	Greensboro, NC	-0.007	City	Texas	-0.030	Non-City
Portland, OR	0.026	City	Los Angeles, CA	-0.008	City	Montana	-0.031	Non-City
Houston, TX	0.024	City	Wyoming	-0.009	Non-City	California	-0.031	Non-City
Milwaukee, WI	0.024	City	Kansas	-0.012	Non-City	Indiana	-0.031	Non-City
Pittsburgh, PA	0.020	City	North Carolina	-0.012	Non-City	Idaho	-0.033	Non-City
Detroit, MI	0.020	City	Florida	-0.013	Non-City	South Dakota	-0.037	Non-City
Rochester, NY	0.019	City	Alabama	-0.014	Non-City	Georgia	-0.037	Non-City
Cleveland, OH	0.019	City	Louisiana	-0.014	Non-City	Arkansas	-0.038	Non-City
New York, NY	0.018	City	West Virginia	-0.015	Non-City	Missouri	-0.040	Non-City
St Louis, MO	0.017	City	Maine	-0.016	Non-City	Oregon	-0.044	Non-City
Columbus, OH	0.017	City	Buffalo, NY	-0.016	City	Minnesota	-0.049	Non-City
San Diego, CA	0.016	City	Ohio	-0.017	Non-City	Nevada	-0.074	Non-City

Ranking occupational complexity

Region	Occ Index	Type	Region	Occ Index	Type	Region	Occ Index	Type
Washington, DC	0.093	City	Virginia Beach, VA	0.013	City	Nebraska	-0.017	Non-City
San Jose, CA	0.093	City	Cincinnati, OH	0.012	City	South Carolina	-0.017	Non-City
Boston, MA	0.070	City	Massachusetts	0.011	Non-City	Michigan	-0.018	Non-City
San Francisco, CA	0.056	City	New Orleans, LA	0.011	City	Illinois	-0.018	Non-City
Seattle, WA	0.056	City	Utah	0.008	Non-City	Riverside, CA	-0.019	City
Denver, CO	0.053	City	Colorado	0.008	Non-City	Maryland	-0.019	Non-City
Baltimore, MD	0.050	City	Rhode Island	0.007	Non-City	Iowa	-0.020	Non-City
Connecticut	0.047	Non-City	Tampa, FL	0.006	City	Miami, FL	-0.020	City
Minneapolis, MN	0.041	City	Vermont	0.004	Non-City	Tennessee	-0.021	Non-City
Atlanta, GA	0.039	City	Arizona	0.002	Non-City	Kentucky	-0.023	Non-City
Philadelphia, PA	0.037	City	New Mexico	0.001	Non-City	Pennsylvania	-0.025	Non-City
Kansas City, MO	0.032	City	New York	0.000	Non-City	North Dakota	-0.025	Non-City
Indianapolis, IN	0.029	City	Oklahoma	-0.003	Non-City	Wisconsin	-0.027	Non-City
New Hampshire	0.029	Non-City	Delaware	-0.003	Non-City	Mississippi	-0.029	Non-City
Dallas, TX	0.029	City	Virginia	-0.004	Non-City	Washington	-0.029	Non-City
Chicago, IL	0.027	City	Greensboro, NC	-0.007	City	Texas	-0.030	Non-City
Portland, OR	0.026	City	Los Angeles, CA	-0.008	City	Montana	-0.031	Non-City
Houston, TX	0.024	City	Wyoming	-0.009	Non-City	California	-0.031	Non-City
Milwaukee, WI	0.024	City	Kansas	-0.012	Non-City	Indiana	-0.031	Non-City
Pittsburgh, PA	0.020	City	North Carolina	-0.012	Non-City	Idaho	-0.033	Non-City
Detroit, MI	0.020	City	Florida	-0.013	Non-City	South Dakota	-0.037	Non-City
Rochester, NY	0.019	City	Alabama	-0.014	Non-City	Georgia	-0.037	Non-City
Cleveland, OH	0.019	City	Louisiana	-0.014	Non-City	Arkansas	-0.038	Non-City
New York, NY	0.018	City	West Virginia	-0.015	Non-City	Missouri	-0.040	Non-City
St Louis, MO	0.017	City	Maine	-0.016	Non-City	Oregon	-0.044	Non-City
Columbus, OH	0.017	City	Buffalo, NY	-0.016	City	Minnesota	-0.049	Non-City
San Diego, CA	0.016	City	Ohio	-0.017	Non-City	Nevada	-0.074	Non-City

Single Index Model: supply (1)

- Adding a subscript i for the individual

$$w_{ir} = \omega_r + \beta_r h_i + \varepsilon_{ir}$$
$$h_i \equiv \widehat{h}_i + \underline{h}_i$$
$$\widehat{h}_i \equiv \omega' x_i$$

  - x_i: vector of standard HC variables (zero mean across r)
  - $\widehat{h}_i, \underline{h}_i$: observed,unobserved component of h_i
  - $\Rightarrow$ nation wide mean normalized to zero
  - ε_{ir}: measurement error in wages
  - ω_r: regional fixed effect
  - β_r: regional return to HC
  - ω: relative contributions of components of x_i to h_i
- **Multiplicative restriction**
  - NLLS estimation (simple repeated OLS)

Single Index Model: supply (1)

- Adding a subscript i for the individual

$$w_{ir} = \omega_r + \beta_r h_i + \varepsilon_{ir}$$
$$h_i \equiv \widehat{h}_i + \underline{h}_i$$
$$\widehat{h}_i \equiv \omega' x_i$$

  - x_i: vector of standard HC variables (zero mean across r)
  - $\widehat{h}_i, \underline{h}_i$: observed,unobserved component of h_i
  - $\Rightarrow$ nation wide mean normalized to zero
  - ε_{ir}: measurement error in wages
  - ω_r: regional fixed effect
  - β_r: regional return to HC
  - ω: relative contributions of components of x_i to h_i

- **Multiplicative restriction**
  - NLLS estimation (simple repeated OLS)

Single Index Model: supply (1)

- Adding a subscript i for the individual

$$w_{ir} = \omega_r + \beta_r h_i + \varepsilon_{ir}$$
$$h_i \equiv \widehat{h}_i + \underline{h}_i$$
$$\widehat{h}_i \equiv \omega' x_i$$

  - x_i: vector of standard HC variables (zero mean across r)
  - $\widehat{h}_i, \underline{h}_i$: observed,unobserved component of h_i
  - $\Rightarrow$ nation wide mean normalized to zero
  - ε_{ir}: measurement error in wages
  - ω_r: regional fixed effect
  - β_r: regional return to HC
  - ω: relative contributions of components of x_i to h_i

- **Multiplicative restriction**
  - NLLS estimation (simple repeated OLS)

Single Index Model: supply (1)

- Adding a subscript i for the individual

$$w_{ir} = \omega_r + \beta_r h_i + \varepsilon_{ir}$$
$$h_i \equiv \widehat{h}_i + \underline{h}_i$$
$$\widehat{h}_i \equiv \omega' x_i$$

  - x_i: vector of standard HC variables (zero mean across r)
  - $\widehat{h}_i, \underline{h}_i$: observed,unobserved component of h_i
  - $\Rightarrow$ nation wide mean normalized to zero
  - ε_{ir}: measurement error in wages
  - ω_r: regional fixed effect
  - β_r: regional return to HC
  - ω: relative contributions of components of x_i to h_i

- **Multiplicative restriction**
  - NLLS estimation (simple repeated OLS)

Single Index Model: supply (2)

- Definitions

$$R_h^2 \equiv \frac{\mathrm{Var}\left[\widehat{h}_i\right]}{\mathrm{Var}\left[\widehat{h}_i\right] + \mathrm{Var}\left[\underline{h}_i\right]}$$

$$\widehat{H}_r \equiv \mathrm{E}\left[\widehat{h}_{ir}|r\right]$$

- Estimation yields estimate for $\widehat{\omega}_r$, not ω_r

$$\widehat{\omega}_r = \omega_r + \beta_r \mathrm{E}\left[\underline{h}_i|r\right]$$

Single Index Model: supply (2)

- Definitions

$$R_h^2 \equiv \frac{\mathrm{Var}\left[\widehat{h}_i\right]}{\mathrm{Var}\left[\widehat{h}_i\right] + \mathrm{Var}\left[\underline{h}_i\right]}$$

$$\widehat{H}_r \equiv \mathrm{E}\left[\widehat{h}_{ir}|r\right]$$

- Estimation yields estimate for $\widehat{\omega}_r$, not ω_r

$$\widehat{\omega}_r = \omega_r + \beta_r \mathrm{E}\left[\underline{h}_i|r\right]$$

Single Index Model: demand

- $$
\begin{aligned}
w_{ir} &= \chi_r + \zeta_r o_i + \varepsilon_{ir} \\
o_i &\equiv \widehat{o}_i + \underline{o}_i \\
\widehat{o}_i &\equiv \zeta' z_i \\
\mathrm{E}_r[\chi_r] = 0, &\qquad \mathrm{E}_r[\zeta_r] = 1 \\
\widehat{O}_r &\equiv \mathrm{E}[\widehat{o}_i | r], \qquad R_o^2 \equiv \ldots
\end{aligned}
$$
    - z_i : a vector of occupations (3 digits = 300 occ)
    - ζ : mean log relative wage in occupation
    - otherwise identical to supply
- **Positive assortative matching**

Single Index Model: demand

- $$\begin{aligned} \color{red}{w_{ir}} &\color{red}{= \chi_r + \zeta_r o_i + \varepsilon_{ir}} \\ o_i &\equiv \widehat{o}_i + \underline{o}_i \\ \widehat{o}_i &\equiv \zeta' z_i \\ \mathrm{E}_r[\chi_r] = 0, &\qquad \mathrm{E}_r[\zeta_r] = 1 \\ \widehat{O}_r &\equiv \mathrm{E}[\widehat{o}_i | r], \qquad R_o^2 \equiv \ldots \end{aligned}$$
  - z_i : a vector of occupations (3 digits = 300 occ)
  - ζ : mean log relative wage in occupation
  - otherwise identical to supply
- **Positive assortative matching**

Single Index Model: demand

- $$
\begin{aligned}
w_{ir} &= \chi_r + \zeta_r o_i + \varepsilon_{ir} \\
o_i &\equiv \widehat{o}_i + \underline{o}_i \\
\widehat{o}_i &\equiv \zeta' z_i \\
\mathrm{E}_r[\chi_r] = 0, &\qquad \mathrm{E}_r[\zeta_r] = 1 \\
\widehat{O}_r \equiv \mathrm{E}[\widehat{o}_i | r], &\qquad R_o^2 \equiv \ldots
\end{aligned}
$$
    - z_i : a vector of occupations (3 digits = 300 occ)
    - ζ : mean log relative wage in occupation
    - otherwise identical to supply
- **Positive assortative matching**

Single Index Model: demand

-

$$
\begin{aligned}
w_{ir} &= \chi_r + \zeta_r o_i + \varepsilon_{ir} \\
o_i &\equiv \widehat{o}_i + \underline{o}_i \\
\widehat{o}_i &\equiv \zeta' z_i \\
\mathrm{E}_r[\chi_r] = 0, &\qquad \mathrm{E}_r[\zeta_r] = 1 \\
\widehat{O}_r &\equiv \mathrm{E}[\widehat{o}_i | r], \qquad R_o^2 \equiv \ldots
\end{aligned}
$$

  - z_i : a vector of occupations (3 digits = 300 occ)
  - ζ : mean log relative wage in occupation
  - otherwise identical to supply
- **Positive assortative matching**

Single Index Model: assignment

- Taking expectations within region r

$$O_r = \frac{\omega_r - \chi_r}{\zeta_r} + \frac{\beta_r}{\zeta_r} H_r$$

$$\beta_r = 1 - \gamma\,(h - o) = 1 - \gamma\left(\frac{\omega_r - \chi_r}{\zeta_r} + \frac{\beta_r - \zeta_r}{\zeta_r} h\right)$$

- **Equal Return Assumption**
    - Autor, Levy & Murnane (2003), Autor & Dorn (2013)

$$\beta_r < \zeta_r \Rightarrow \frac{\mathrm{d}^2 w}{\mathrm{d}h^2} > 0 : \text{ polarization}$$

$$\beta_r = \zeta_r : \text{ linearity}$$

$$\mathrm{Cor}\left[\beta_r, \zeta_r\right] \overset{\text{empirical}}{=} 0.86$$

$$\beta_r = 1 - \gamma\left(\frac{\omega_r - \chi_r}{\zeta_r}\right)$$

Single Index Model: assignment

- Taking expectations within region r

$$O_r = \frac{\omega_r - \chi_r}{\zeta_r} + \frac{\beta_r}{\zeta_r} H_r$$

$$\beta_r = 1 - \gamma \left(h - o \right) = 1 - \gamma \left(\frac{\omega_r - \chi_r}{\zeta_r} + \frac{\beta_r - \zeta_r}{\zeta_r} h \right)$$

- **Equal Return Assumption**
  - Autor, Levy & Murnane (2003), Autor & Dorn (2013)

$$\beta_r < \zeta_r \Rightarrow \frac{\mathrm{d}^2 w}{\mathrm{d} h^2} > 0 : \text{ polarization}$$

$$\beta_r = \zeta_r : \text{ linearity}$$

$$\mathrm{Cor}\left[\beta_r, \zeta_r\right] \overset{\text{empirical}}{=} 0.86$$

$$\beta_r = 1 - \gamma \left(\frac{\omega_r - \chi_r}{\zeta_r} \right)$$

Single Index Model: Proportionality Assumption

- Assumption (dropping i again)

$$\mathrm{E}\left[\widehat{o}|h,r\right] = R_o^2 \mathrm{E}\left[o|h,r\right]$$
$$\mathrm{E}\left[\underline{o}|h,r\right] = \left(1-R_o^2\right)\mathrm{E}\left[o|h,r\right]$$

  - decomposition of h uninformative on decompostion o
- $E = \mathrm{Var}[w]/\mathrm{Var}[\varepsilon] = 0.30$: Angrist & Krueger (1999)

Table : Intraregional Variance Decomposition

No.	Variance	Data	Formula	Calculated
1.	$\mathrm{Cov}\left[\widehat{h},\widehat{o}\right]/\mathrm{Var}[w|r]$	21%	$(1-E)\,R_h^2 R_o^2$	20%
2.	$\mathrm{Cov}\left[\widehat{h},\underline{o}\right]/\mathrm{Var}[w|r]$	16%	$(1-E)\,R_h^2\left(1-R_o^2\right)$	17%
3.	$\mathrm{Cov}[\underline{h},\widehat{o}]/\mathrm{Var}[w|r]$	16%	$(1-E)\left(1-R_h^2\right)R_o^2$	17%
4.	$\mathrm{Cov}[\underline{h},\underline{o}]/\mathrm{Var}[w|r]$	17%	$(1-E)\left(1-R_h^2\right)\left(1-R_o^2\right)$	15%
5.	$\mathrm{Var}[e]/\mathrm{Var}[w|r]$	30%	E	30%

Single Index Model: Proportionality Assumption

- Assumption (dropping i again)

$$\mathrm{E}\left[\widehat{o}|h,r\right] = R_o^2 \mathrm{E}\left[o|h,r\right]$$
$$\mathrm{E}\left[\underline{o}|h,r\right] = \left(1 - R_o^2\right) \mathrm{E}\left[o|h,r\right]$$

  - decomposition of h uninformative on decompostion o

- $E = \mathrm{Var}[w] / \mathrm{Var}[\varepsilon] = 0.30$: Angrist & Krueger (1999)

Table : Intraregional Variance Decomposition

No.	Variance	Data	Formula	Calculated
1.	$\mathrm{Cov}\left[\widehat{h}, \widehat{o}\right] / \mathrm{Var}[w\|r]$	21%	$(1-E)\, R_h^2 R_o^2$	20%
2.	$\mathrm{Cov}\left[\widehat{h}, \underline{o}\right] / \mathrm{Var}[w\|r]$	16%	$(1-E)\, R_h^2 \left(1-R_o^2\right)$	17%
3.	$\mathrm{Cov}[\underline{h}, \widehat{o}] / \mathrm{Var}[w\|r]$	16%	$(1-E)\left(1-R_h^2\right) R_o^2$	17%
4.	$\mathrm{Cov}[\underline{h}, \underline{o}] / \mathrm{Var}[w\|r]$	17%	$(1-E)\left(1-R_h^2\right)\left(1-R_o^2\right)$	15%
5.	$\mathrm{Var}[e] / \mathrm{Var}[w\|r]$	30%	E	30%

Single Index Model: Proportionality Assumption

- Assumption (dropping i again)

$$\mathrm{E}\left[\widehat{o}|h,r\right] = R_o^2 \mathrm{E}\left[o|h,r\right]$$
$$\mathrm{E}\left[\underline{o}|h,r\right] = \left(1 - R_o^2\right) \mathrm{E}\left[o|h,r\right]$$

  - decomposition of h uninformative on decompostion o

- $E = \mathrm{Var}[w] / \mathrm{Var}[\varepsilon] = 0.30$: Angrist & Krueger (1999)

Table : Intraregional Variance Decomposition

No.	Variance	Data	Formula	Calculated
1.	$\mathrm{Cov}\left[\widehat{h},\widehat{o}\right]/\mathrm{Var}[w|r]$	21%	$(1-E)\,R_h^2 R_o^2$	20%
2.	$\mathrm{Cov}\left[\widehat{h},\underline{o}\right]/\mathrm{Var}[w|r]$	16%	$(1-E)\,R_h^2\left(1-R_o^2\right)$	17%
3.	$\mathrm{Cov}[\underline{h},\widehat{o}]/\mathrm{Var}[w|r]$	16%	$(1-E)\left(1-R_h^2\right)R_o^2$	17%
4.	$\mathrm{Cov}[\underline{h},\underline{o}]/\mathrm{Var}[w|r]$	17%	$(1-E)\left(1-R_h^2\right)\left(1-R_o^2\right)$	15%
5.	$\mathrm{Var}[e]/\mathrm{Var}[w|r]$	30%	E	30%

Single Index Model: testing proportionality

- Regression on $\widehat{h}$ and $\underline{h} + \varepsilon$ separately

$$\widehat{h} = \chi^{(1)\prime} z + \varepsilon^{(1)}$$

$$R^2 = R_o^2 R_h^2 = 0.28 \overset{\text{empircally}}{=} 0.36$$

$$\underline{h} + \varepsilon = w_i - \omega_r - \widehat{h} = \chi^{(2)\prime} z + \varepsilon^{(2)}$$

$$(1 - E)\left(1 - R_h^2\right) R_o^2 = 0.17 \overset{\text{empircally}}{=} 0.12$$

$$\text{Cor}\left[\chi^{(1)\prime} z, \chi^{(2)\prime} z\right] \overset{\text{empircally}}{=} 0.70$$

- Correction $\widehat{\omega}_r$

$$H_r = R_h^{-2} \widehat{H}_r = 1.89 \widehat{H}_r$$

$$\omega_r = \widehat{\omega}_r - \frac{1 - R_h^2}{R_h^2} \widehat{H}_r = \widehat{\omega}_r - 0.89 \widehat{H}_r$$

Single Index Model: testing proportionality

- Regression on $\widehat{h}$ and $\underline{h} + \varepsilon$ separately

$$\widehat{h} = \chi^{(1)\prime} z + \varepsilon^{(1)}$$

$$R^2 = R_o^2 R_h^2 = 0.28 \overset{\text{empircally}}{=} 0.36$$

$$\underline{h} + \varepsilon = w_i - \omega_r - \widehat{h} = \chi^{(2)\prime} z + \varepsilon^{(2)}$$

$$(1 - E)\left(1 - R_h^2\right) R_o^2 = 0.17 \overset{\text{empircally}}{=} 0.12$$

$$\text{Cor}\left[\chi^{(1)\prime} z, \chi^{(2)\prime} z\right] \overset{\text{empircally}}{=} 0.70$$

- Correction $\widehat{\omega}_r$

$$H_r = R_h^{-2} \widehat{H}_r = 1.89 \widehat{H}_r$$

$$\omega_r = \widehat{\omega}_r - \frac{1 - R_h^2}{R_h^2} \widehat{H}_r = \widehat{\omega}_r - 0.89 \widehat{H}_r$$

Single Index Model: testing proportionality

- Regression on $\widehat{h}$ and $\underline{h}+\varepsilon$ separately

$$\widehat{h} = \chi^{(1)\prime} z + \varepsilon^{(1)}$$

$$R^2 = R_o^2 R_h^2 = 0.28 \overset{\text{empircally}}{=} 0.36$$

$$\underline{h} + \varepsilon = w_i - \omega_r - \widehat{h} = \chi^{(2)\prime} z + \varepsilon^{(2)}$$

$$(1 - E)\left(1 - R_h^2\right) R_o^2 = 0.17 \overset{\text{empircally}}{=} 0.12$$

$$\text{Cor}\left[\chi^{(1)\prime} z, \chi^{(2)\prime} z\right] \overset{\text{empircally}}{=} 0.70$$

- Correction $\widehat{\omega}_r$

$$H_r = R_h^{-2} \widehat{H}_r = 1.89 \widehat{H}_r$$

$$\omega_r = \widehat{\omega}_r - \frac{1 - R_h^2}{R_h^2} \widehat{H}_r = \widehat{\omega}_r - 0.89 \widehat{H}_r$$

Single Index Model: testing proportionality

- Regression on $\widehat{h}$ and $\underline{h} + \varepsilon$ separately

$$\widehat{h} = \chi^{(1)\prime} z + \varepsilon^{(1)}$$

$$R^2 = R_o^2 R_h^2 = 0.28 \overset{\text{empircally}}{=} 0.36$$

$$\underline{h} + \varepsilon = w_i - \omega_r - \widehat{h} = \chi^{(2)\prime} z + \varepsilon^{(2)}$$

$$(1 - E)\left(1 - R_h^2\right) R_o^2 = 0.17 \overset{\text{empircally}}{=} 0.12$$

$$\text{Cor}\left[\chi^{(1)\prime} z, \chi^{(2)\prime} z\right] \overset{\text{empircally}}{=} 0.70$$

- Correction $\widehat{\omega}_r$

$$H_r = R_h^{-2} \widehat{H}_r = 1.89 \widehat{H}_r$$

$$\omega_r = \widehat{\omega}_r - \frac{1 - R_h^2}{R_h^2} \widehat{H}_r = \widehat{\omega}_r - 0.89 \widehat{H}_r$$

Single Index Model: testing proportionality

- Regression on $\widehat{h}$ and $\underline{h} + \varepsilon$ separately

$$\widehat{h} = \chi^{(1)\prime} z + \varepsilon^{(1)}$$

$$R^2 = R_o^2 R_h^2 = 0.28 \overset{\text{empircally}}{=} 0.36$$

$$\underline{h} + \varepsilon = w_i - \omega_r - \widehat{h} = \chi^{(2)\prime} z + \varepsilon^{(2)}$$

$$(1 - E)\left(1 - R_h^2\right) R_o^2 = 0.17 \overset{\text{empircally}}{=} 0.12$$

$$\text{Cor}\left[\chi^{(1)\prime} z, \chi^{(2)\prime} z\right] \overset{\text{empircally}}{=} 0.70$$

- Correction $\widehat{\omega}_r$

$$H_r = R_h^{-2} \widehat{H}_r = 1.89 \widehat{H}_r$$

$$\omega_r = \widehat{\omega}_r - \frac{1 - R_h^2}{R_h^2} \widehat{H}_r = \widehat{\omega}_r - 0.89 \widehat{H}_r$$

Conclusions

1. Large interregional wage differentials and returns to HC
2. Simple variables explain 80% interregional wage variation
3. High HC demand activities concentrated in cities
4. Strong agglomeration externalities
5. ... driven by demand for HC (**not** supply!)
6. Demand for HC clusters more strongly than supply
7. ... despite the fact that housing is a normal good
8. Cities must create a lot of amenities to generate this outcome