Wage Inequality and Cities

Winter School on Inequality and Social Welfare Theory

Nathaniel Baum-Snow

University of Toronto

Rotman School of Management

Goals

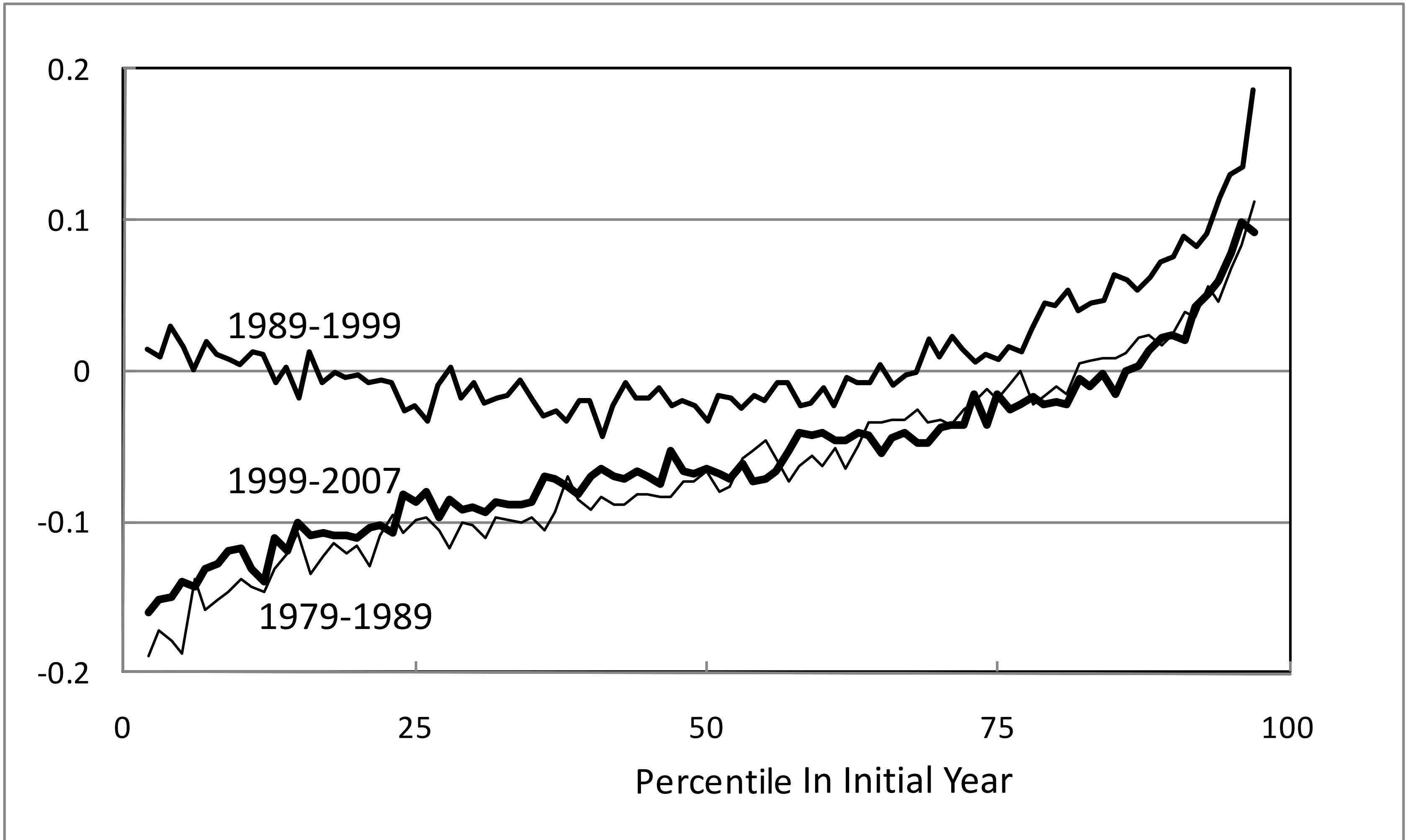
- Highlight some important facts in the data that demonstrate a potentially important role for local labor markets in understanding trends in wage and income inequality
    - Average wage gaps between cities of different sizes
    - Differences in wage distributions between cities of different sizes
- Rationalize these facts in a spatial equilibrium environment
- Examine evidence on potential mechanisms driving these differences
    - Labor & macro literatures
    - Nature of agglomeration economies
- Specify what we don't know yet
    - A lot!

U.S. Nationwide Growth in Wage Inequality

Log wage gaps between “skilled” and “unskilled”

Skilled Worker Definition Unskilled Worker Definition	Some College+ High School-	College+ Some College-	College+ High School-	College Only High School Only
Panel A: All Workers				
1980	0.30	0.47	0.50	0.38
1990	0.38	0.54	0.62	0.46
2000	0.43	0.57	0.66	0.50
2005-7	0.48	0.62	0.72	0.55
Panel B: Manufacturing Workers Only				
1980	0.36	0.56	0.60	0.48
1990	0.42	0.59	0.67	0.53
2000	0.46	0.63	0.72	0.57
2005-7	0.53	0.70	0.80	0.64

U.S. Nationwide Growth in Wage Inequality



U.S. Nationwide Growth in Wage Inequality

Year	Variance	Total 90-50 Gap	50-10 Gap
1979	0.21	0.53	0.63
1989	0.29	0.62	0.70
1999	0.34	0.72	0.67
2004-7	0.39	0.81	0.74
79-07 Change	0.18	0.28	0.11

$$\ln w_{idst} = \alpha_{dst} + \varepsilon_{idst}$$

Year	Total Variance	Between Variance	Residual Variance
1979	0.21	0.05	0.16
1989	0.29	0.08	0.21
1999	0.34	0.10	0.25
2004-7	0.39	0.12	0.27
79 to 07 Change	0.18	0.07	0.11

Separating Out Components Between versus Within Local Labor Markets

- The urban wage premium is an important part of the “between” component

Mean Log Wage Relative to Small Cities and Rural Areas
Decennial Census 5% PUMS

	1980	**1990**	**2000**
MSAs: 250,000 - 1.5 million	0.14***	0.18***	0.19***
	(0.01)	(0.01)	(0.01)
MSAs: > 1.5 million	0.23***	0.31***	0.32***
	(0.01)	(0.01)	(0.01)
R-squared	0.03	0.05	0.04

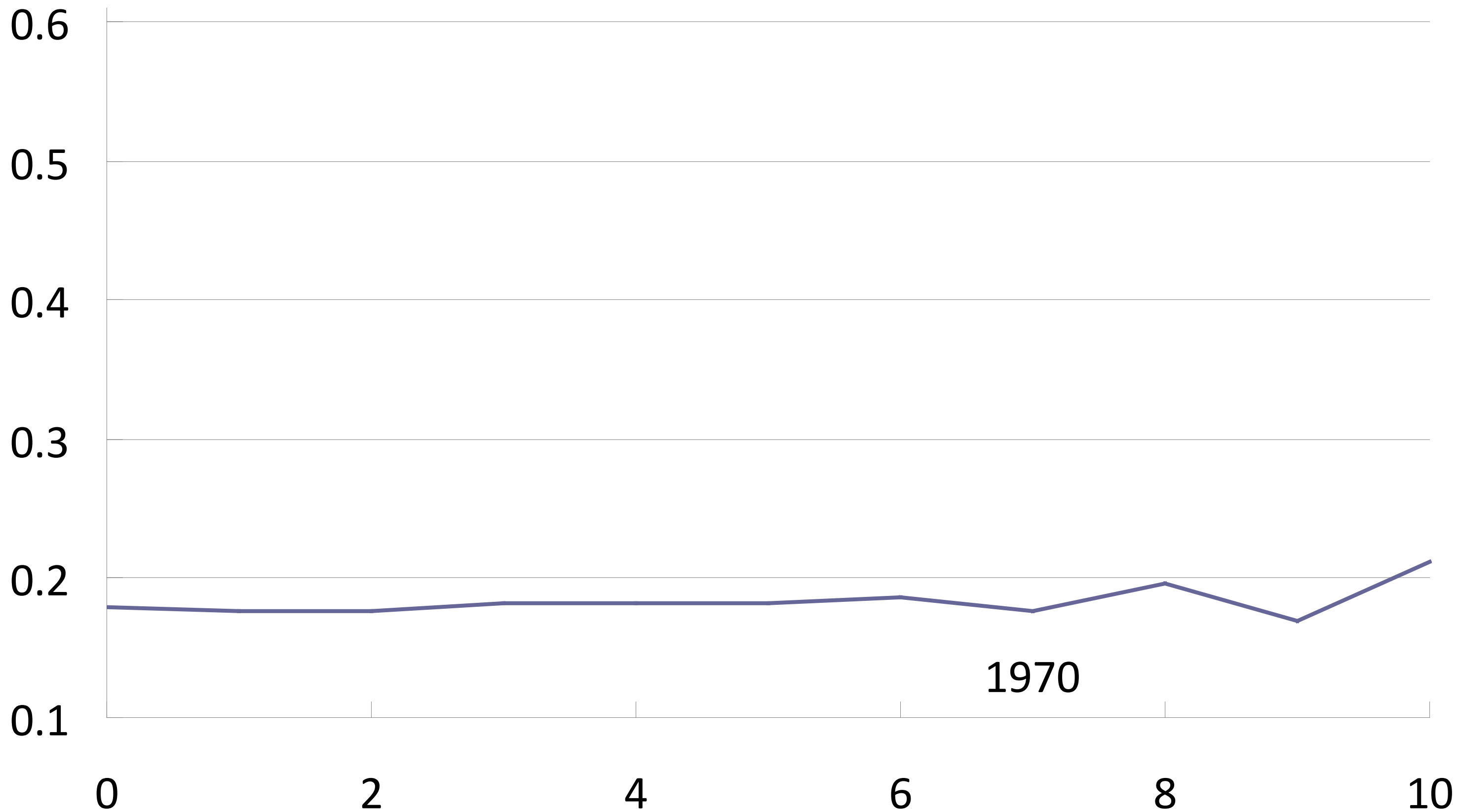
- There is a lot of work on understanding mechanisms driving the urban wage premium
  - Baum-Snow & Pavan (2012), De La Roca & Puga (2015)

- There is less research about why it has increased over time

Variance of Weekly Wages by City Size

0 = Rural Areas

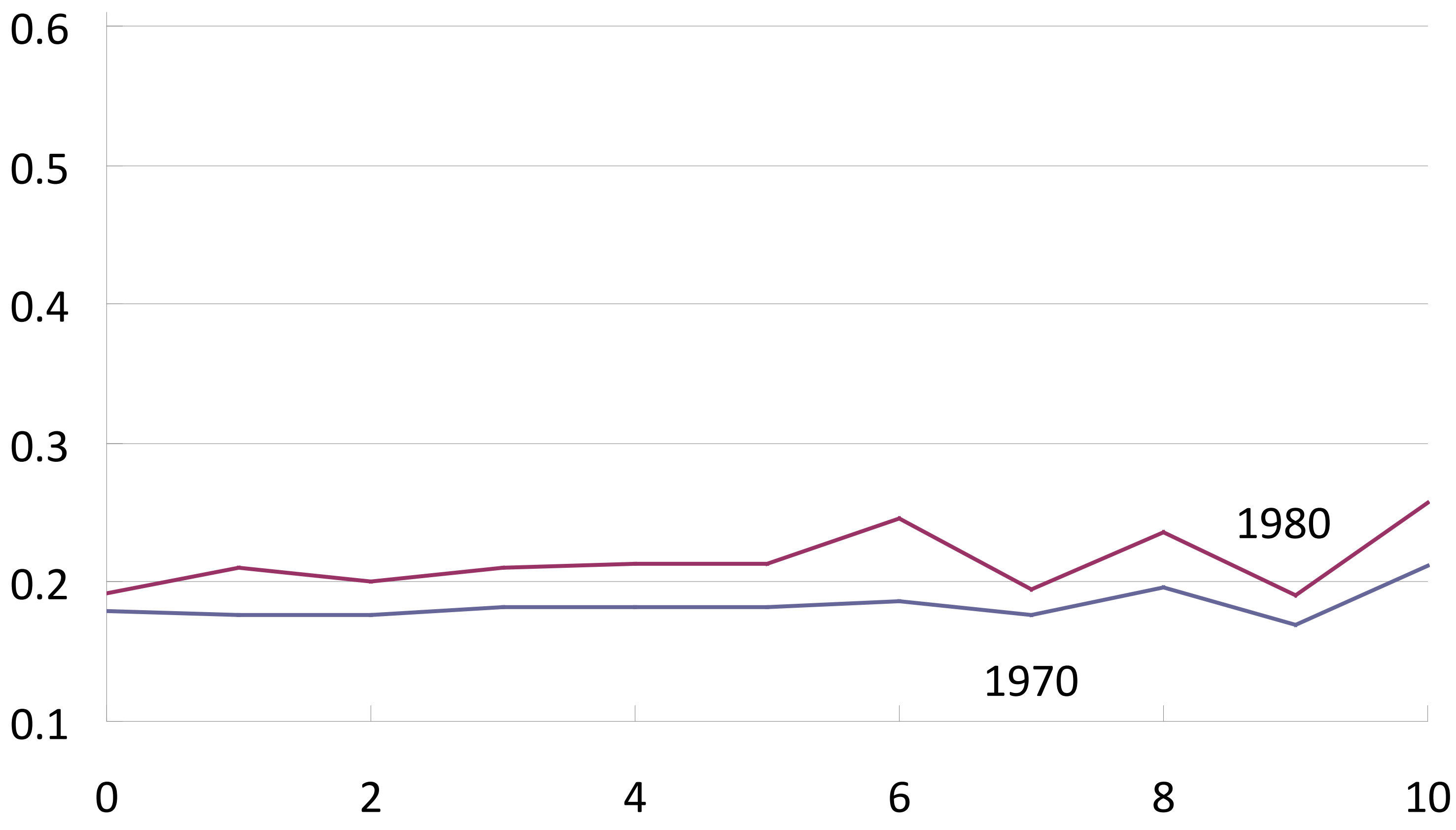
1-10 = Deciles of Urban Population by City Size



Variance of Weekly Wages by City Size

0 = Rural Areas

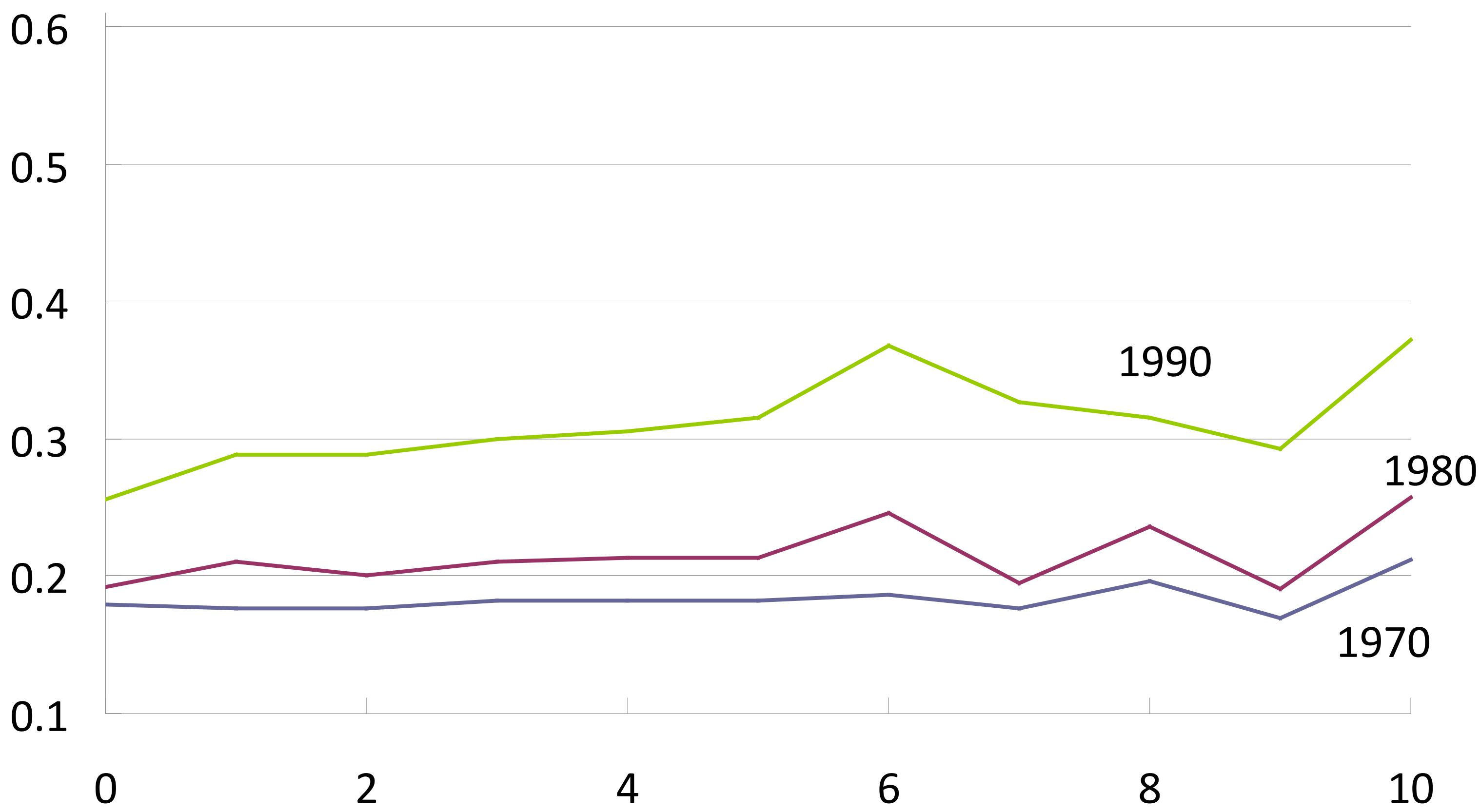
1-10 = Deciles of Urban Population by City Size



Variance of Weekly Wages by City Size

0 = Rural Areas

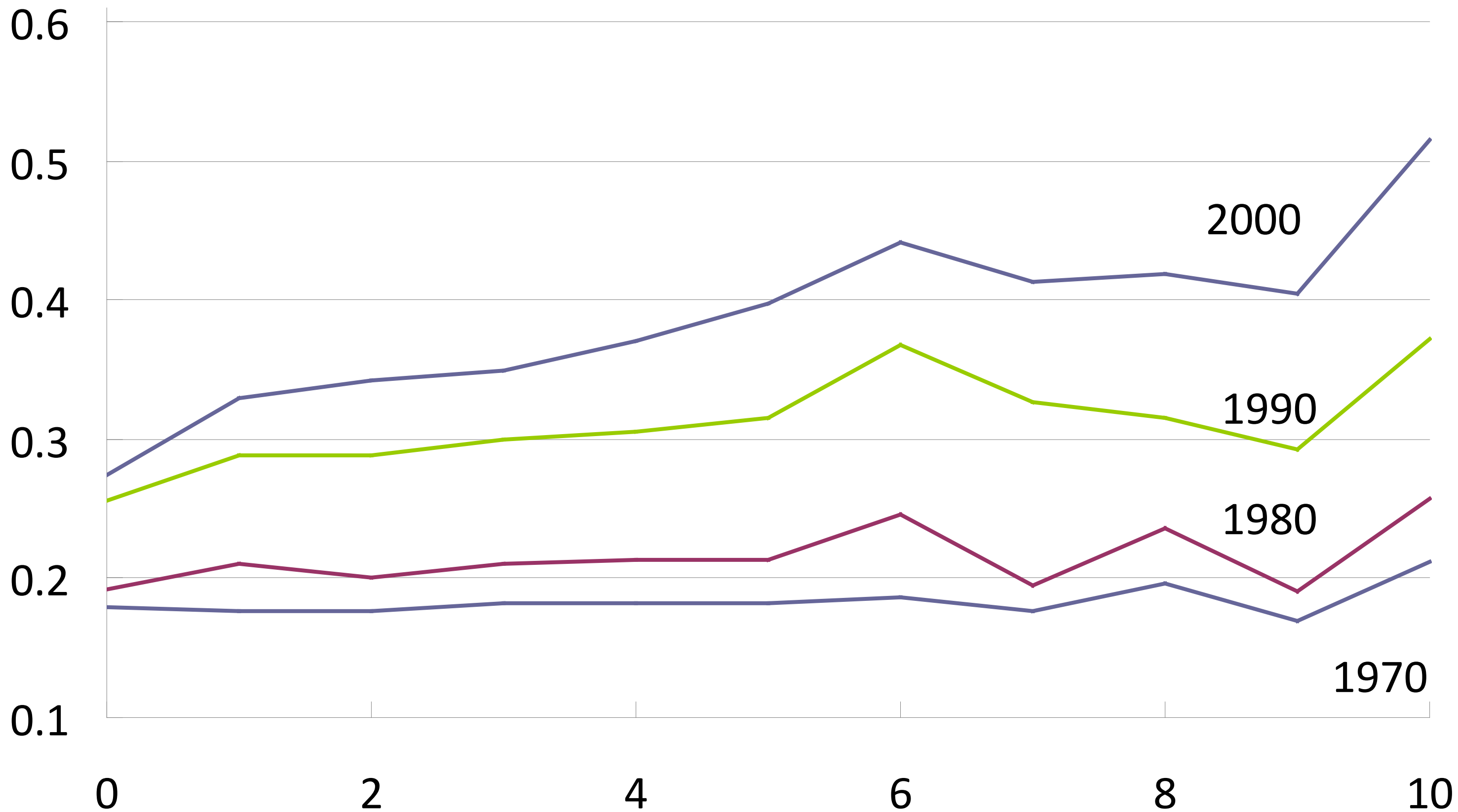
1-10 = Deciles of Urban Population by City Size



Variance of Weekly Wages by City Size

0 = Rural Areas

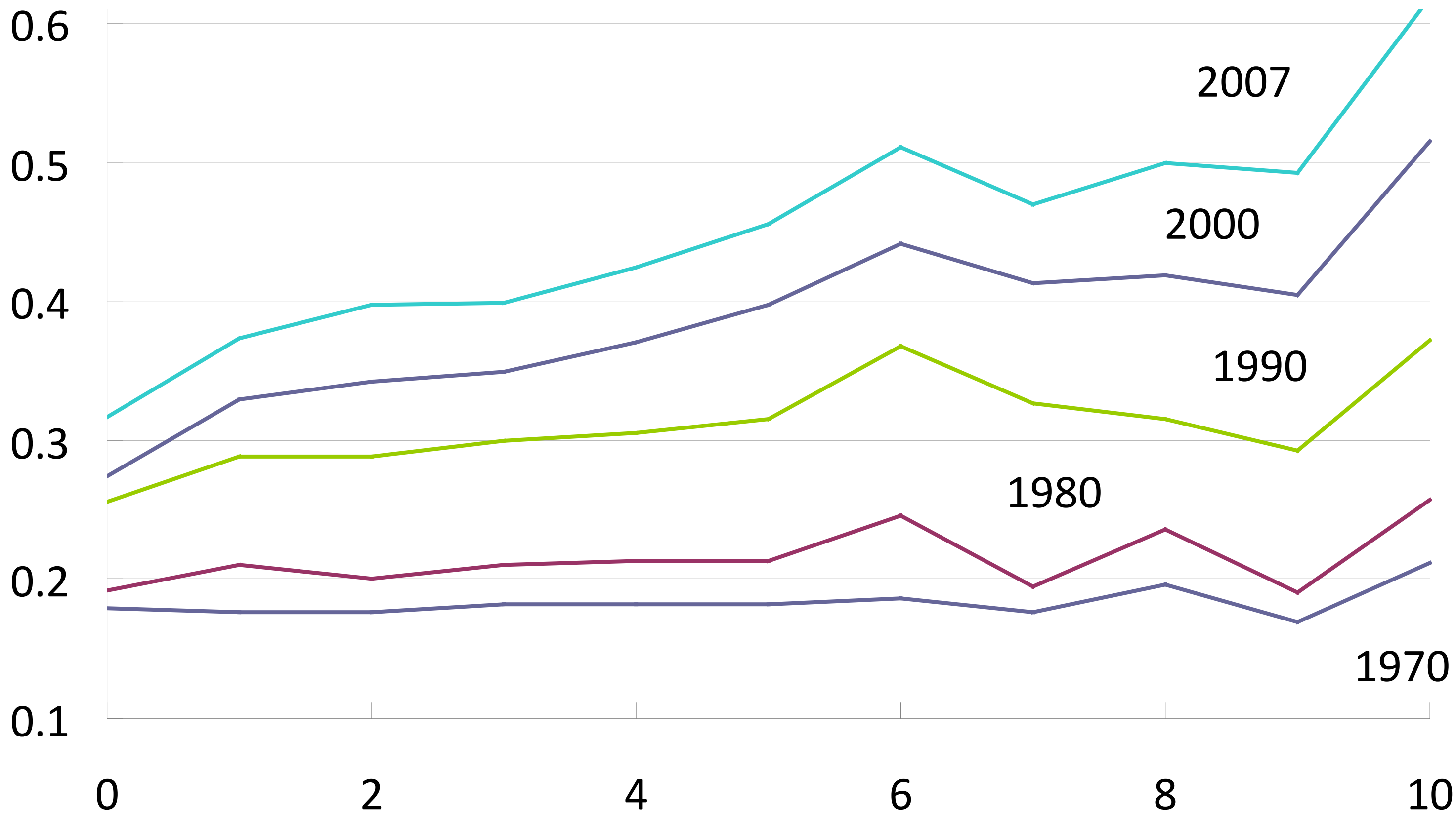
1-10 = Deciles of Urban Population by City Size



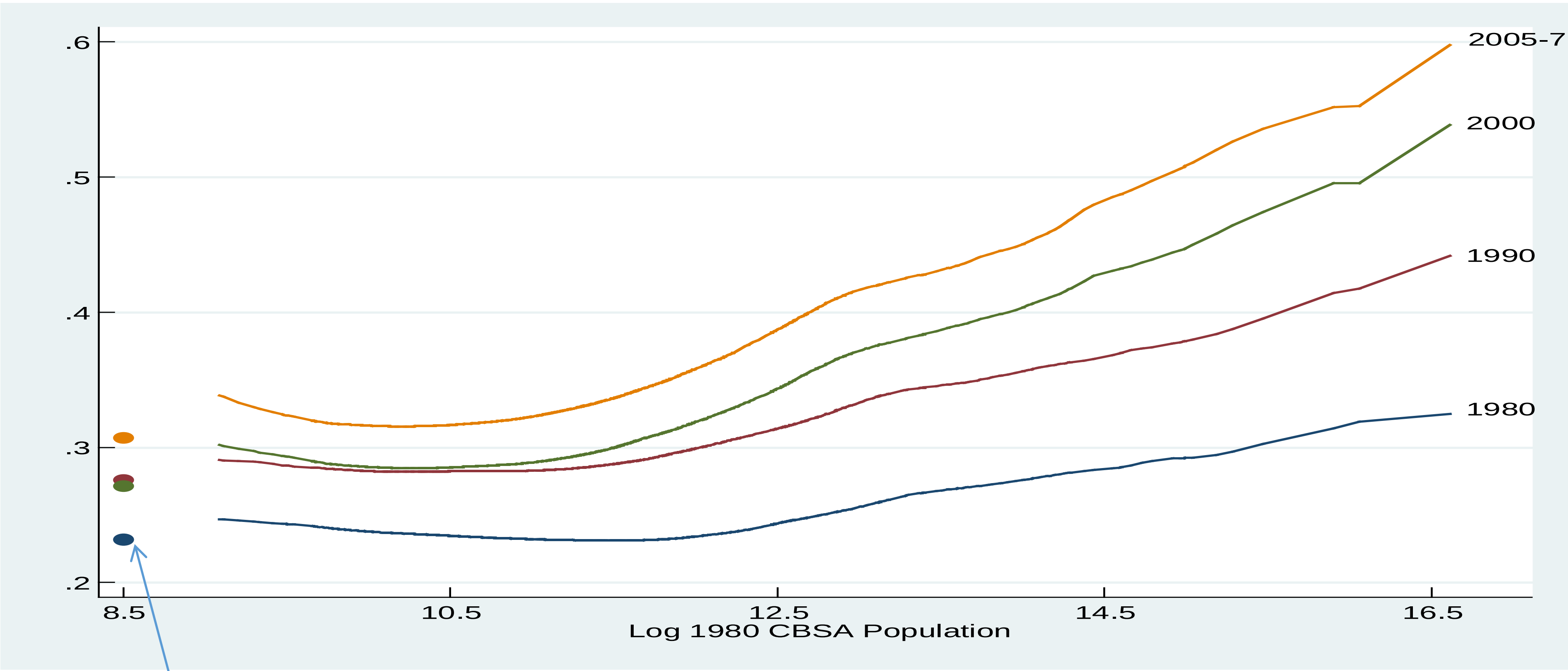
Variance of Weekly Wages by City Size

0 = Rural Areas

1-10 = Deciles of Urban Population by City Size

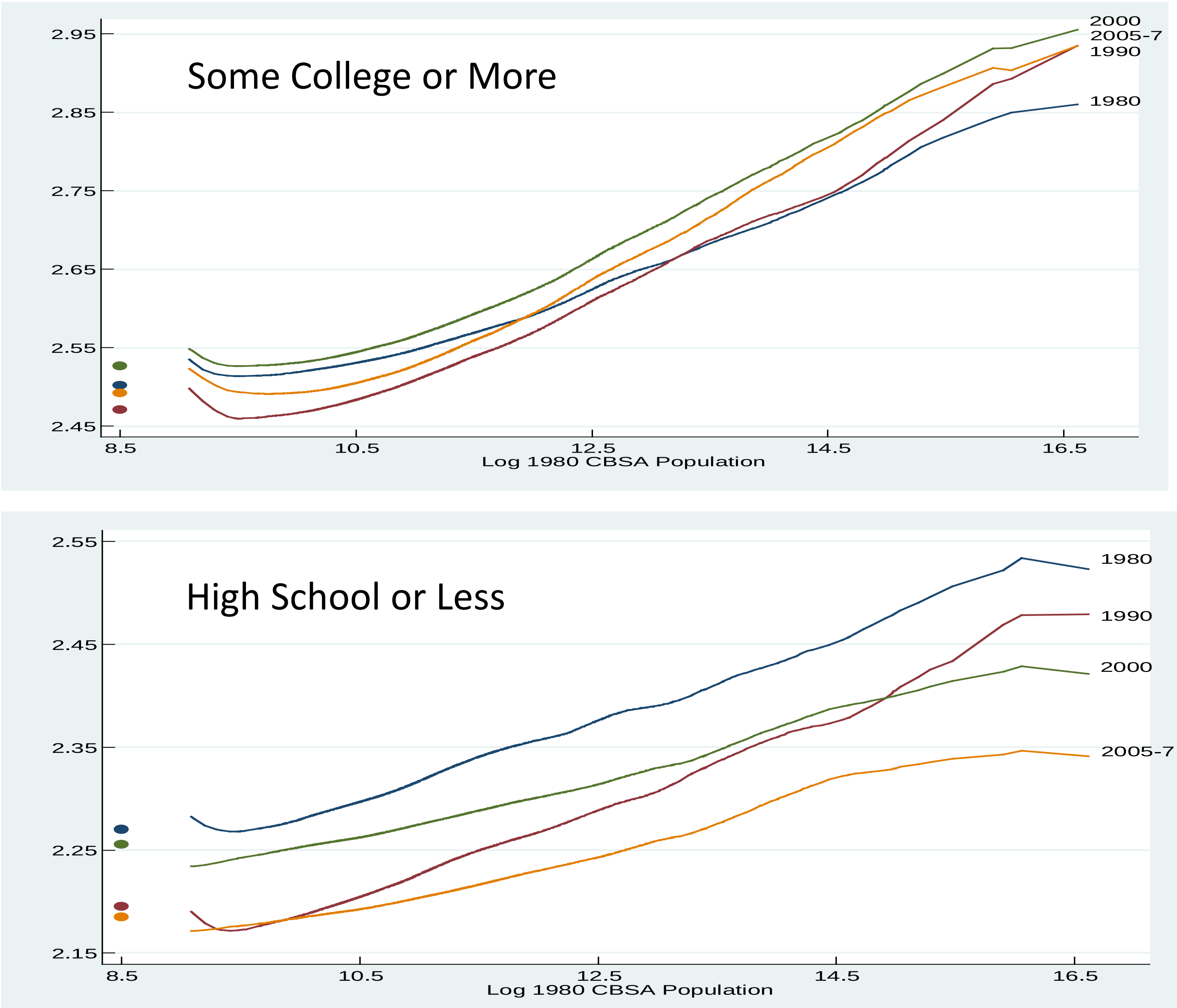


Some College+ vs. High School- Wage Gaps by City Size
All Workers

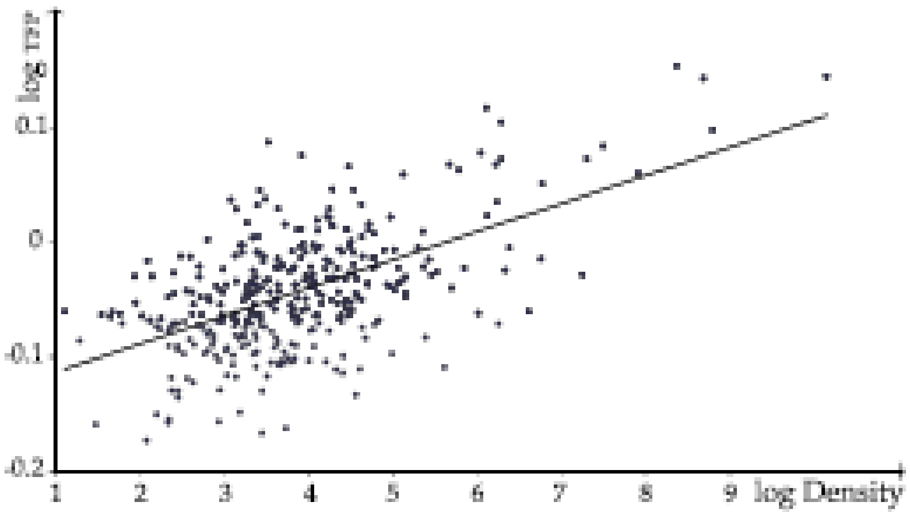


Rural Areas

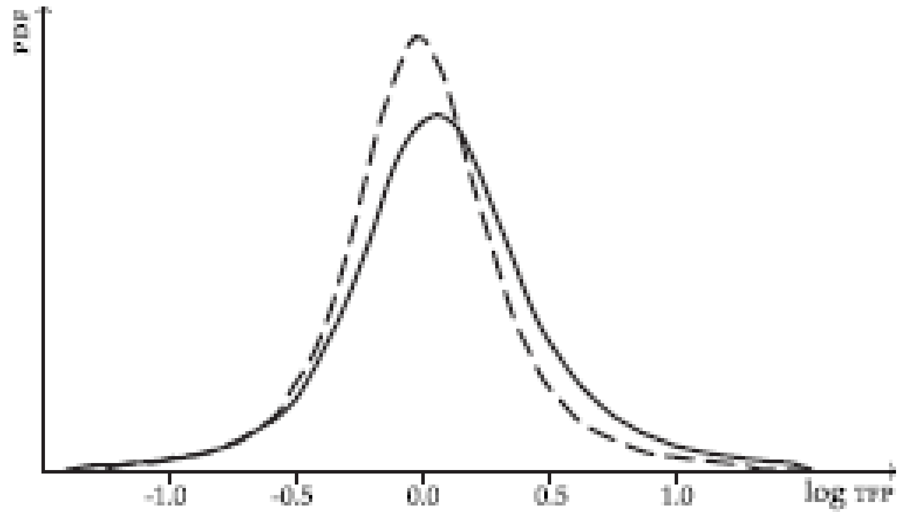
Skilled & Unskilled Wages by City Size – All Workers



Firm TFP in France
Combes et al. (2012)



Panel (a)
The relationship between mean log TFP and log density for French employment areas



Panel (b)
Distribution of log TFP for all sectors, employment areas above (solid) versus below (dashed) median density

FIGURE 1.—The productive advantages of large cities.

Rationalizing Cross-Sectional Wage Patterns in the Data

"Rosen-Roback" long-run spatial equilibrium conditions

- Producers of tradeables indifference condition
    - Lets us learn about traded productivity differences across locations
- Consumer indifference condition
    - Lets us learn about quality of life differences across locations
    - Allows us to think about the labor supply environment

Indifference Condition for Firms Producing Tradeables

Suppose firms are in long-run location equilibrium, produce tradeables using a CRS technology, are perfectly mobile and have labor (L), capital (K) and local goods (R) as inputs. From this, we see that wage gaps across locations are closely related to productivity differences across these locations

$$\overline{\pi} = \max_{L,K,R} \{ a_j F(L,K,R) - w_j L - rK - p_j R \}$$

$$\textit{implies}$$

$$\ln w_j - \ln w_{j'} = \frac{\ln a_j - \ln a_{j'}}{\phi_L} - \frac{\phi_R}{\phi_L}[\ln p_j - \ln p_{j'}]$$

- ϕ gives factor input shares

- Calibrations of parameters indicate that nominal city size wage gaps overstate productivity differences by about 15 percent
    - Implication is that wage differences across locations are closely related to productivity differences across locations

- Additional labor factors of production do not change this conclusion much since all skill types have higher wages in larger cities
    - More on skill heterogeneity below

What Factors Could Account For The City Size Wage Gap?

1. Sorting of more able workers into larger cities

2. Agglomeration economies (Duranton and Puga, 2004):

   - “Sharing” :
     - Indivisibilities in intermediate inputs
     - Risk sharing
     - Input market pooling

   Empirically amounts to differences in wage intercepts across location types

   - “Learning” :
     - Different rates of human capital accumulation by workers

   Differences in returns to experience across location types

   - “Matching” :
     - Thicker labor markets mean more rapid ascending of job ladders
     - Higher variance distributions of firm-worker match quality exist in thicker labor markets

   Differences in job turnover and firm-worker component of wages across location types

- We can imagine many of these mechanisms may be likely to be skill-biased

Long-Run Consumer Indifference Conditions

- Indirect utilities across equated across locations j for those of each type c for each skill group g

$$V(p_j, w_j^g; q_j^{gc}) = \overline{v}^{gc}$$

local prices

wages

local amenities

- Generates the following equilibrium conditions across locations:

$$d\ln w^g = \beta_g d\ln p - \frac{\partial \ln V / \partial \ln q}{\partial \ln V / \partial \ln w} d\ln q^{gc}$$

expenditure share on local goods

U is typically specified as qu(x,H)
u() homothetic, making this a constant

- There are important lifecycle considerations for consumers though that turn out to matter a lot in practice

Using Consumer Indifference to Recover Labor Supply Functions to Cities

- Following Notowidigdo (2013), suppose there is a migration friction such that each individual draws from a distribution of migration costs to each location such that the last (highest cost) migrant of type gc to j pays M(.) in migration cost.
  - This generates a wedge in differences, assuming constant consumer amenities over time.
  - (Diamond (2016) generates upward sloping labor supply with taste shocks.)
- Differentiating over time, we get the following law of motion, imposing

$$d\ln {w_j}^g - \beta_g d\ln p_j - M^g(d\ln_j N^{gc}, {N_{j0}}^{gc}) = (1/\frac{d\ln V}{d\ln w})d\ln \bar{v}^{gc}$$

Constant for many standard
Preference specifications

Deriving Labor Supply Conditions

• Solving for the change in indirect utility and specifying $M^{g} = \alpha_{g}^{1} d\ln N_{j}^{gc} - \alpha_{g}^{2} N_{j0}^{gc}$

$$\begin{aligned} d\ln N_j^g \;=\; & \frac{1}{\alpha_g^1}\left[d\ln w_j^g - \beta_g d\ln p_j\right] - \frac{1}{\alpha_g^1}\sum_c \frac{N_j^{gc}}{N_j^g}\sum_k \frac{N_k^{gc}}{N^{gc}}\left[d\ln w_k^g - \beta_g d\ln p_k + \alpha_g^2 N_{k0}^{gc}\right] \\ & + \sum_c \frac{N_j^{gc}}{N_j^g}\left[d\ln N^{gc} + \frac{\alpha_g^2}{\alpha_g^1} N_{j0}^{gc}\right] \end{aligned}$$

own effect

competition effect

immigration effect (shock)

• The existence of migration costs means that locals will not full respond to real wage pressure that occurs because of lots of immigrant arrivals

• We will use this later on to achieve exogenous variation in skilled and unskilled labor quantities across local labor markets

City Costs

- On one side of the ledger were the productivity advantages of density, on the other side is city costs, or what generates differences in p_j across cities

- The little empirical evidence we have (Combes, Duranton & Gobillon, 2014) indicates that $dlnp_j/dlnN_j$ is comparable to $dlnw_j/dlnN_j$, both at about 0.04

- Urban costs are typically derived using a monocentric model of city structure

- Welfare depends on both wages and costs

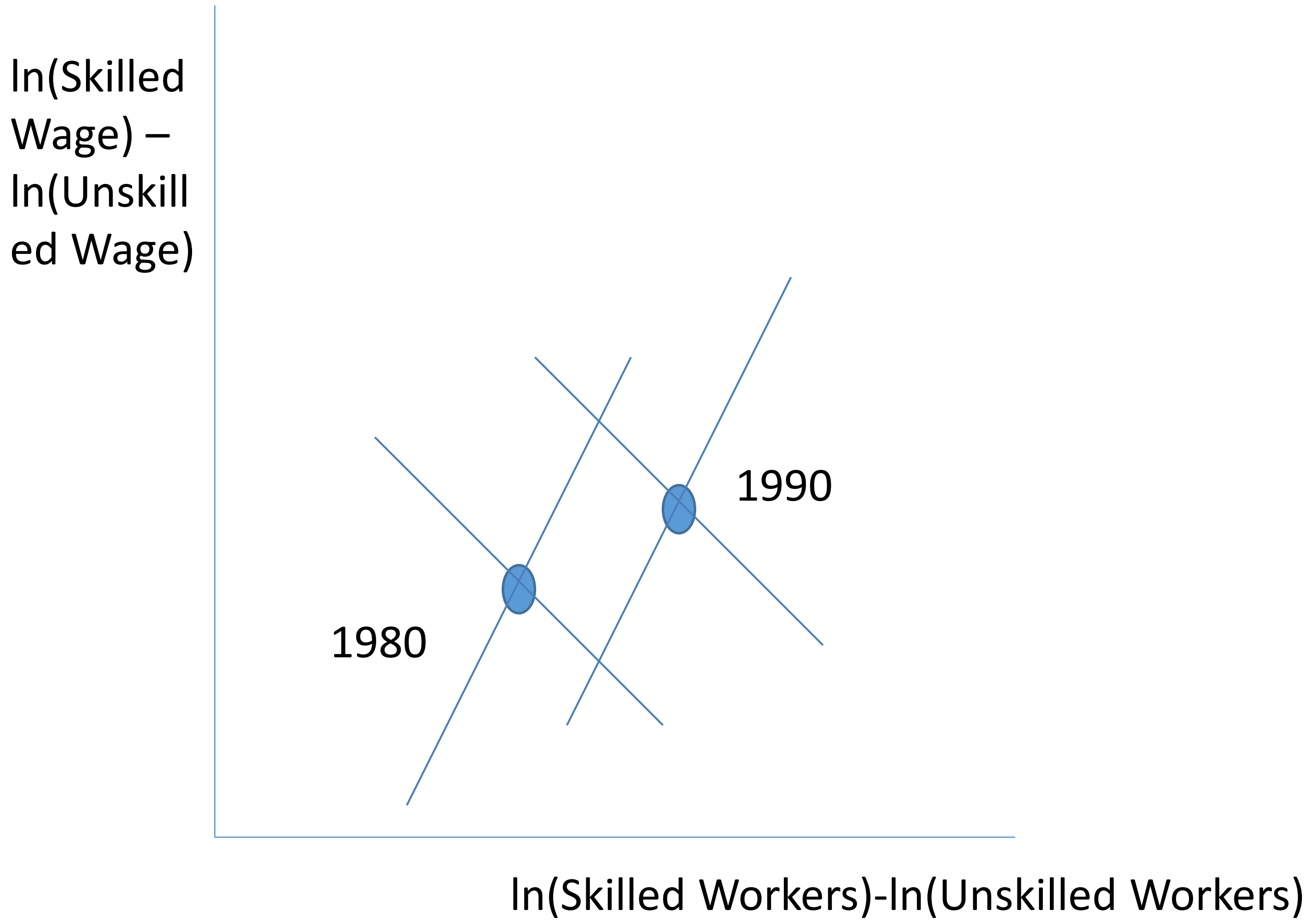
    - Moretti's (2013) evidence that "real wage inequality" has not increased as much as nominal wage inequality

    - The increased dispersion in marginal products of labor across locations and skills motivates focusing mostly on trying to understand causal channels that operate through labor demand

Investigating City Size and Inequality Over Time

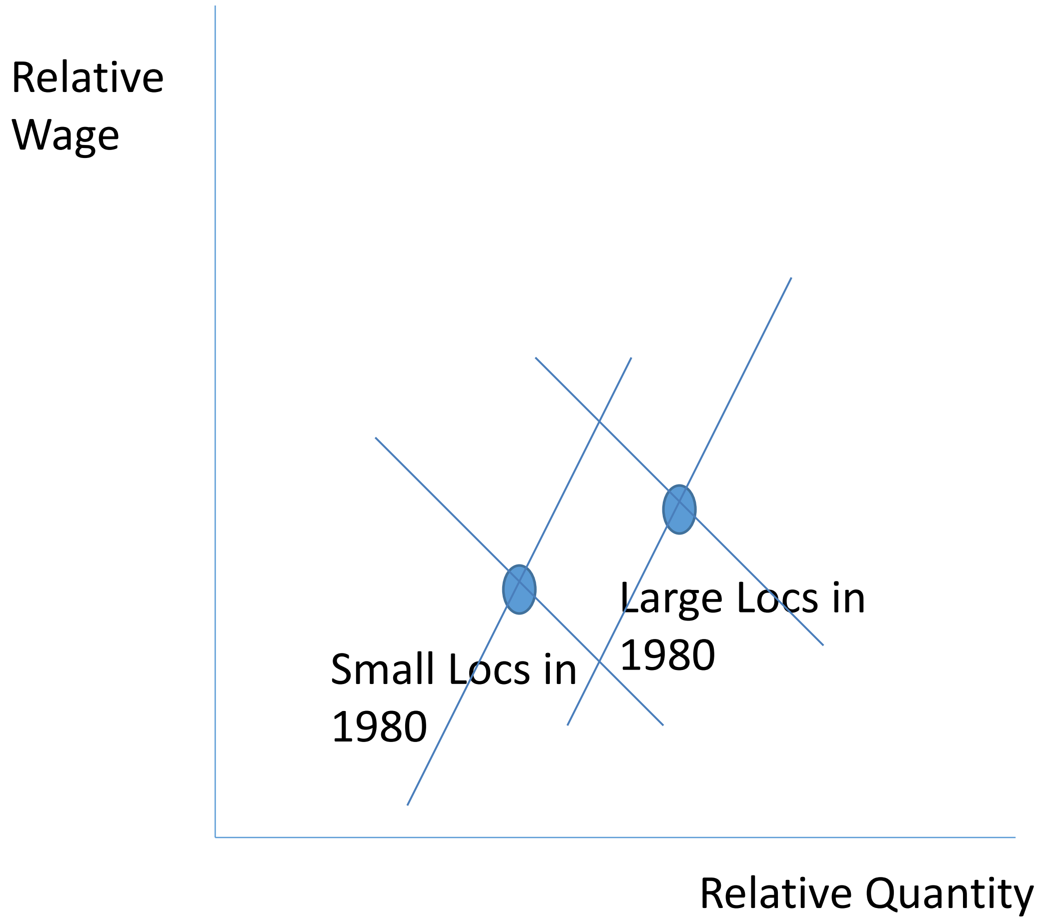
- What role does “city size” have in causing the observed increases in “between” and “residual” wage dispersion?
    - How much is from city size’s effects on prices vs. quantities of observed and unobserved skills?

Katz & Murphy’s (1992) Seminal Observation



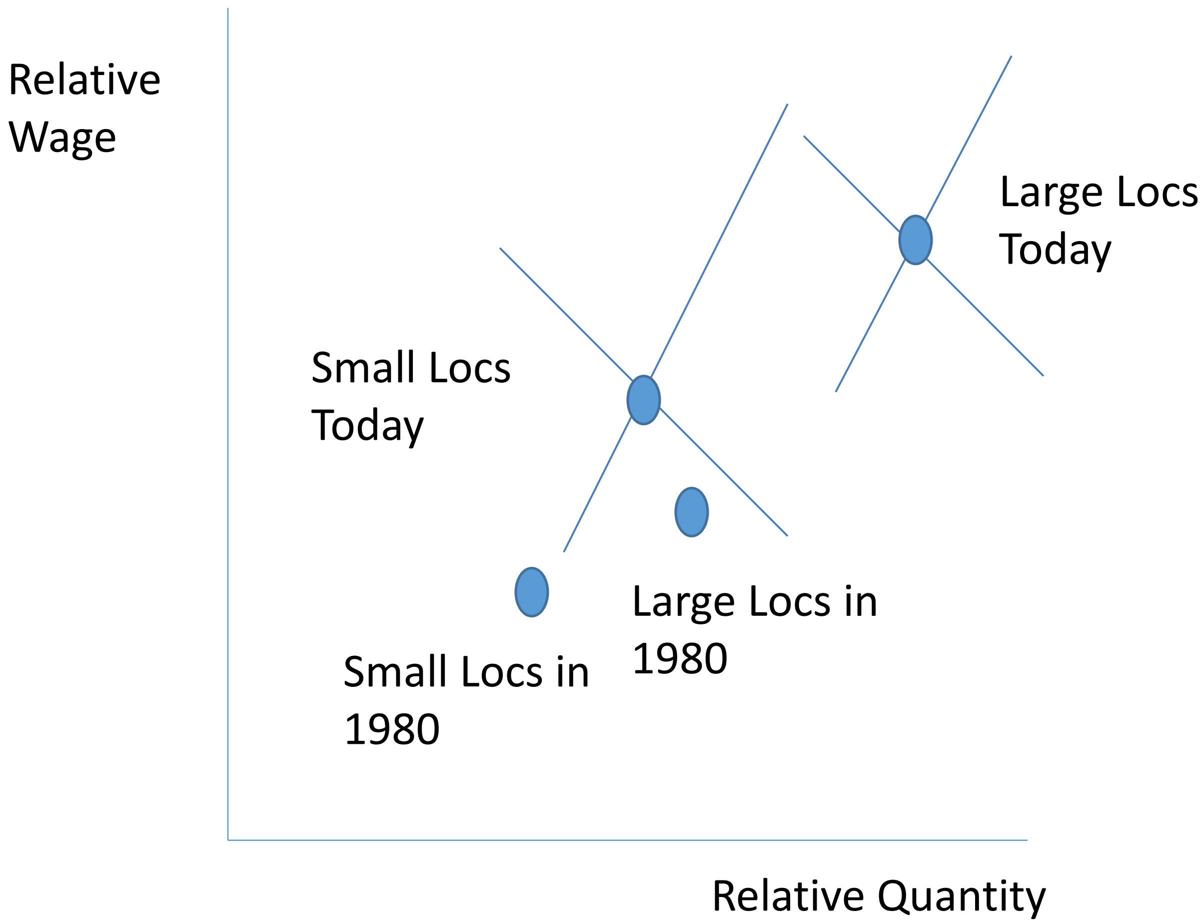
- Increase in relative quantities and relative wages means that there must have been a labor demand shift
- Write down a model to quantify this shift and account for growth inwage inequality

Applying the Same Logic to Cities



Relative Relative Demand Shifts

- Larger relative demand gaps across locations today than in 1980



- Strong evidence of relative relative demand shifts as in Katz & Murphy (1992)

Framework for Examining The Roles of Demographics and City Size (Baum-Snow & Pavan, 2013)

- As shown by DiNardo, Fortin & Lemieux (1996), the distribution of wages at time t can be broken up into, means m, residual "price" distributions g and "quantity" distributions h

$$f_t(\varepsilon) = \int g_t(\varepsilon \mid x, s) h_t(x, s) \, ds dx$$

$$a_t(w) = \int g_t(w - m_t(x, s) \mid x, s) h_t(x, s) \, ds dx$$

ε = residual

$x =$ demographics (75 ageXeduc cells)

$s =$ city size (0 - 10)

m = demographic group mean

Accounting for Quantities Only

- Replace the distribution of people within each demographic group across city size categories with that from 1980. This gives us a counterfactual distribution that takes out changes in sorting on observed skill across locations, but allows the nationwide shifts in the composition of the population that occurred to remain.

$$f_t^{\,q}(\varepsilon)=\int g_t(\varepsilon \mid x, s)h_{a1979}(s \mid x)h_{bt}(x)dsdx$$

Additionally Accounting for Prices

- Within each skill group, maintain the same percentile relationships between rural locations and each urban size category that existed in 1980. Take the evolution of wage inequality in rural locations as exogenous.

  - This is an implementation of the "Changes in Changes" model of Athey and Imbens (2006) where rural locations are the "control" group and each city size category is a different "treatment" group

  - Murphy, Juhn & Pierce (1993) do something similar using national distributions

  - Suppose that each residual is the product of a price and quantity of unobserved skill, and the distributions of unobserved skill do not change over time

$$\varepsilon_t(x,s) = \rho_t(x,s)u(\phi \mid x,s) \text{ implies}$$

$$\varepsilon_{\mathrm{t}}{}^{p}(x,s) = \rho_{1979}(x,s)\frac{\rho_t(x,0)}{\rho_{1979}(x,0)}u(\phi \mid x,s)$$

Accounting for Prices and Quantities

- Calculate the counterfactual price distribution $g_t^p(\varepsilon|x,s)$ by combining all of these individual counterfactual residuals

- Generate counterfactual distribution of residuals taking into account both prices and quantities

$$f_t^c(\varepsilon) = \int g_t^p(\varepsilon \mid x, s) h_{a1979}(s \mid x) h_{bt}(x) ds dx$$

Investigating the Importance of the Level of Sorting on Observed Skill for Price Results

- Carry out the same exercise on the same set of residuals $\varepsilon_t(x,s)$ but place everyone in the same demographic group. This yields the following counterfactual distribution:

$$f_t^{\,n}(\varepsilon) = \int g_t^{\,n}(\varepsilon \mid s) h_{1979}(s) ds$$

- Comparison of inequality measures in the resulting counterfactual distribution to that in the one that fully accounts for prices and quantities reveals the importance of accounting for differences in the composition of observed demographic groups across different locations

Counterfactual Residual Growth – Reductions Relative to Actual

	1	2	3
Calculated Using	$f_t^q(\varepsilon)$	$f_t^c(\varepsilon)$	$f_t^n(\varepsilon)$
X Set	Full Demog	Full Demog	No Demog
Adjustment	Quantities	Residual Prices & Quantities	Residual Prices & Quantities
Panel A: Variance			
1979 to 1989	0%	20%	26%
1979 to 1999	-2%	49%	59%
1979 to 2004-7	-2%	35%	43%
Panel B: 90 - 50 Percentile Gap			
1979 to 1989	0%	21%	46%
1979 to 1999	-1%	36%	62%
1979 to 2004-7	-1%	30%	50%
Panel C: 50 - 10 Percentile Gap			
1979 to 1989	0%	7%	-2%
1979 to 1999	-3%	104%	88%
1979 to 2004-7	-3%	58%	51%

Constructing Counterfactual log Wage Distributions

- Our counterfactual mean wages are very conservative as they allow the overall mean wages to move as they did in equilibrium. If we instead index to rural locations, we get much larger effects.

$$m_t^c(x,s) = \int m_t(x,s)h_{at}(s\,|\,x)ds + [m_{1979}(x,s) - \int m_{1979}(x,s)h_{at}(s\,|\,x)ds]$$

- The resulting counterfactual distribution is:

$$a_t^c(w) = \int g_t^p\left(w - m_t^c(x,s)\,|\,x,s\right)h_{a1979}(s\,|\,x)h_{bt}(x)dsdx$$

- Carry out the same exercise assigning everyone to the same demographic group

Counterfactual log Wage Growth

	1	2	3	4
Calculated Using	$a_t^{\,q}(w)$	$a_t^{\,\varepsilon}(w)$	$a_t^{\,c}(w)$	$a_t^{\,n}(w)$
X Set	Full Demog	Full Demog	Full Demog	No Demog
Adjustment	Quantities	Residual Prices & Quantities	Total Prices & Quantities	Total Prices & Quantities
Panel A: Variance				
1979 to 1989	-2%	10%	16%	26%
1979 to 1999	-2%	31%	34%	50%
1979 to 2004-7	-1%	21%	23%	43%
Panel B: 90 - 50 Percentile Gap				
1979 to 1989	-4%	-14%	-10%	17%
1979 to 1999	2%	14%	17%	50%
1979 to 2004-7	-2%	17%	20%	51%
Panel C: 50 - 10 Percentile Gap				
1979 to 1989	0%	6%	16%	14%
1979 to 1999	-19%	71%	78%	102%
1979 to 2004-7	0%	19%	20%	41%

Evaluating the Role of Industry Composition Shifts

- Data not rich enough to do this totally nonparametrically as in the earlier analysis

- Break up "Between" and "Residual" components using the following regression, now incorporating industries j

$$\ln w_{idjst} = \alpha_{dst} + \beta_{djt} + \delta_{jst} + \varepsilon_{idjst}$$

- Calculate variances as follows, where θ represents share:

$$V_t(\ln w_{idjst}) = \sum_{d,j,s} \theta_{djst} V(\alpha_{dst} + \beta_{djt} + \delta_{jst}) + \sum_{d,j,s} \theta_{djst} V(\varepsilon_{idjst})$$

- Generate counterfactuals analogously to before by replacing the θ and variance components to have 1980 profiles

Evaluating the Role of Industry Composition Shifts

	1	2	3	4	5	6
Calculated Using		$f_t^{\,q}(\varepsilon)$	$a_t^{\,q}(w)$		$f_t^{\,c}(\varepsilon)$	$a_t^{\,c}(w)$
		Quantities			Prices & Quantities	
	Between	Residual	Total	Between	Residual	Total
Panel A: Demographics, Industry and City Size						
1979 to 1989	-2%	0%	-1%	9%	16%	13%
1979 to 1999	0%	-2%	-1%	7%	36%	26%
1979 to 2004-7	2%	-2%	0%	4%	22%	15%
Panel B: Demographics and City Size						
1979 to 1989	-4%	-1%	-2%	8%	21%	15%
1979 to 1999	-1%	-2%	-2%	10%	45%	33%
1979 to 2004-7	1%	-2%	-1%	9%	31%	22%

Conclusions from Baum-Snow & Pavan (2013)

- Something about city size is related to the increase in wage inequality nationwide in the U.S. 1980-2010.
  - Absent the more rapid rise in skill prices in larger cities than smaller cities, nationwide wage inequality would have grown by about 25-33% less
- But what is it?
  - Think about mechanisms

Canonical Models for understanding the Nationwide Rise in Wage Inequality

- Acemoglu & Autor (2011)

$$Y = \left[(A_L L)^{\frac{\sigma-1}{\sigma}} + (A_H H)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

  - Rise in wage inequality interpreted as increases in A_H over time, $\sigma>1$ for substitutes
  - Lots of evidence that high and low skilled labor are substitutes, more on that later

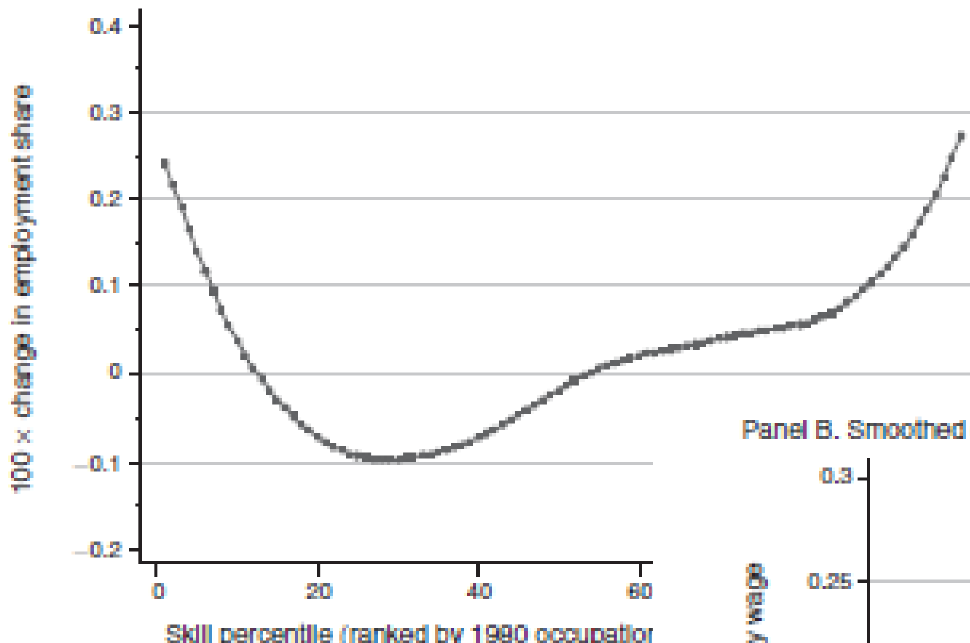
- Krusell et al. (2000 tell the following story: declines in the price of capital + capital-skill complementarity = increases in demand for skill, essentially enriching the model to make the argument they don't need the factor augmenting elements

$$Y_t = A_t K_s^{\alpha} \left[c U_t^{\sigma} + (1-c)[\lambda K_t^{\rho} + (1-\lambda) S_t^{\rho}]^{\sigma/\rho} \right]^{\frac{1-\alpha}{\sigma}}$$

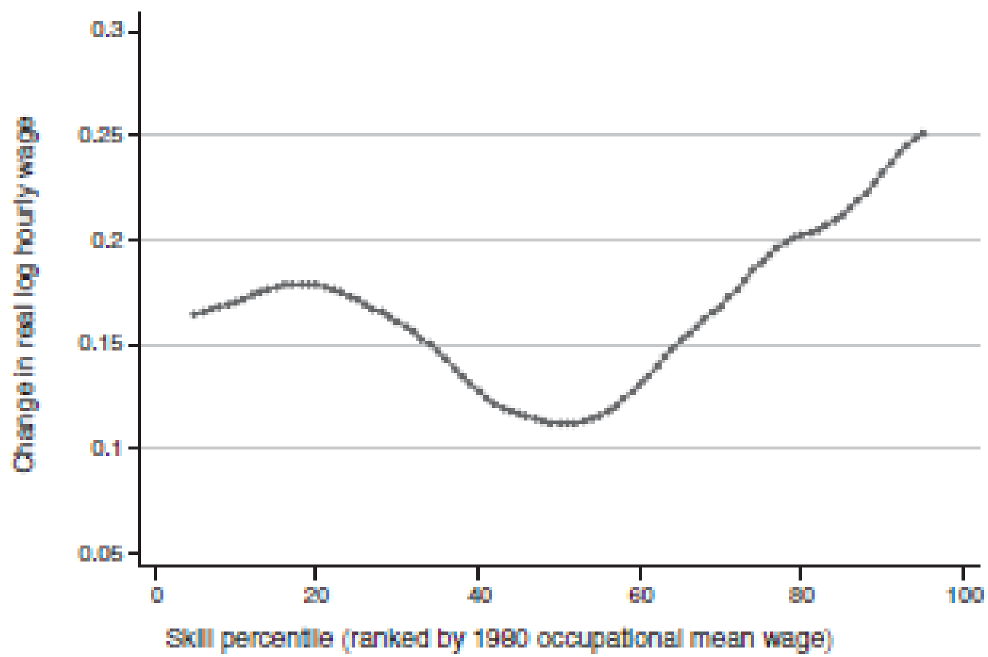
  - If $\sigma > \rho$, there is capital-skill complementarity

Polarization (Autor & Dorn, 2013)

Panel A. Smoothed changes in employment by skill percentile, 1980–2005



Panel B. Smoothed changes in real hourly wages by skill percentile, 1980–2005



Polarization (Autor & Dorn, 2013)

- Goods sector

$$Y_g = L_a^{1-\beta}\left[(\alpha_r L_r)^{\mu} + (\alpha_k K)^{\mu}\right]^{\beta/\mu},$$

$\frac{1}{1-\mu} < 1$ so routine labor and capital are substitues

=> routine labor or capital and abstract labor are relative complements

- Service sector

$$Y_s = \alpha_s L_m$$

- Low skilled workers can either go into routine or manual labor
- CES preferences for some variety between goods and services
- Story of increasing productivity and falling capital price over time
- Use variation across metro areas and Bartik instruments to estimate parameters

Cross-Sectional Evidence

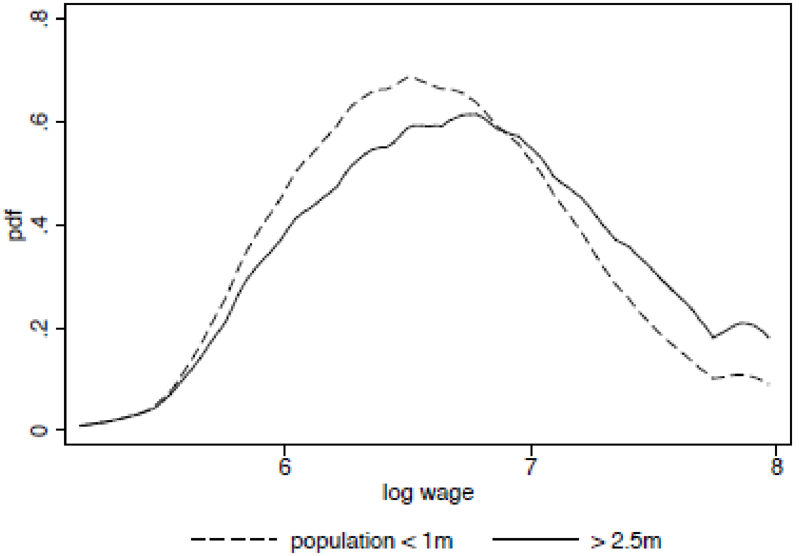
- Beaudry, Doms & Lewis (2010) and Lewis (2011) look at PC and technology adoption as a function of skilled labor inputs

- Ciccone & Peri (2005) generate estimates and summarize myriad other estimates of the elasticity of substitution between skilled and unskilled workers in a two-factor model
    - Between 1.4 and 2 -> more substitutable than Cobb-Douglas production
    - None of these estimates in the literature use variation across local labor markets to identify effects, though Ciccone & Peri (2005) use state/year variation

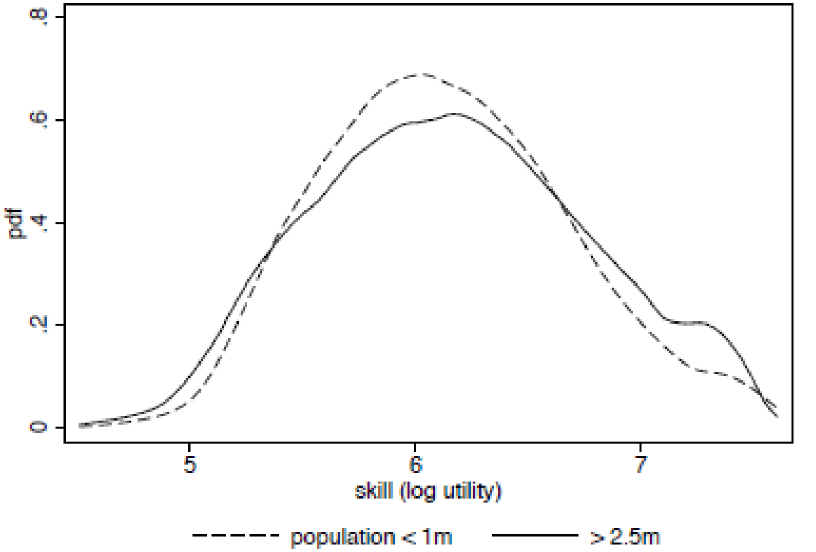
- Burstein, Vogel, Morales (2015) use a CES model with occupations as factors that can incorporate capital-skill complementarity
    - Isomorphism between local labor markets and occupations in their setup

Why Is Wage Inequality Higher in Larger Cities?

- Eeckhout, Pinhiero & Schmidheiny (2014) consider a model that delivers this pattern plus thicker tails of the ability distribution in larger cities



10th percentile: pop < 1m = 5.93, pop > 2.5m = 5.99, diff = 0.065*** (0.007)
90th percentile: pop < 1m = 7.36, pop > 2.5m = 7.56, diff = 0.198*** (0.007)



10th percentile: pop < 1m = 5.44, pop > 2.5m = 5.36, diff = -0.074*** (0.006)
90th percentile: pop < 1m = 6.86, pop > 2.5m = 6.99, diff = 0.132*** (0.009)

- The central feature is "extreme skill complementarity", which has a production technology like:

$$A_j[(m_{1j}{}^{\gamma} y_1 + m_{3j}{}^{\gamma} y_3)^{\lambda} + m_{2j} y_2]$$

$$\lambda > 1 \text{ for } \mathrm{m}_1 \text{ and } \mathrm{m}_3 \text{ complements}$$

Baum-Snow, Freedman, Pavan (2016) Production Technology

- Standard CES production function as in Krusell et al. (2000), with the addition of agglomeration economies which are potentially factor-biased and aggregation to the CBSA rather than national level

- We think of this production function as holding for each point in time, with one observation per CBSA j

$$Y_j = A_j \left[c A_u^{\sigma} D_j^{\sigma \mu_u} U_j^{\sigma} + (1-c)\left(\lambda A_k^{\rho} D_j^{\rho \mu_k} K_j^{\rho} + (1-\lambda) A_s^{\rho} D_j^{\rho \mu_s} S_j^{\rho} \right)^{\frac{\sigma}{\rho}} \right]^{\frac{1}{\sigma}}$$

Output

TFP

CBSA Pop.

Unskilled Labor

Factor-Specific Productivity

Capital

Skilled Labor

- Can think of this production function as being CES in U and a composite input made up of K and S
- Different factor augmenting coefficients on each of them

- We explore the other (less standard) nesting in a robust check

Using This Technology

- Estimate parameters of this production technology using data on capital, skilled labor and unskilled labor aggregated to the metropolitan area level

    - Reduced form analysis points to existence of capital-skill complementarity, as found by Lewis (2011) using different data

    - Structurally estimate a 3-4 equation system to recover production function parameters

    - Identification is achieved by leveraging plausibly exogenous variation in the relative supply of skilled workers from immigration shocks, as in Card (2001) and Lewis (2011)

- Consider relative importance of mechanisms for driving the strengthening relationship between city size and wage gaps

    - Shifts in the factor biases of agglomeration economies
    - Capital-skill complementarity

Empirical Patterns Over Time – Manufacturing Workers

	Manufacturing Workers					
	Raw Counts			Efficiency Units		
	$\Delta\ln(w^S/w^U)$	$\Delta\ln(S/U)$	$\Delta\ln(K/S)$	$\Delta\ln(w^S/w^U)$	$\Delta\ln(S/U)$	$\Delta\ln(K/S)$
log(1980 CBSA Pop)	0.014***	-0.002	0.014***	0.011***	0.001	0.016***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
1990-2000	0.000	-0.335***	0.218***	-0.025***	-0.308***	0.196***
Indicator	(0.003)	(0.008)	(0.011)	(0.003)	(0.008)	(0.011)
2000-2007	0.003	-0.524***	0.268***	-0.000	-0.516***	0.254***
Indicator	(0.003)	(0.008)	(0.011)	(0.003)	(0.008)	(0.011)
Constant	0.031***	0.598***	-0.246***	0.033***	0.591***	-0.261***
	(0.003)	(0.007)	(0.009)	(0.002)	(0.007)	(0.009)
Observations	2,766	2,766	2,202	2,766	2,766	2,202
R-Squared	0.113	0.639	0.256	0.144	0.624	0.230

- Significantly more rapid increases in wage gaps in larger cities
- No significant increase in skill intensity in larger cities
- Significantly more rapid capital intensification in larger cities

Deriving Estimating Equations

- Because our exogenous variation from immigration shocks exists in changes rather than levels, we must derive estimating equations that are functions of dln(S/U)

- We assume that the capital rental rate v is not a function of CBSA j

- Define these shares: $\omega_j^c = \frac{\lambda D_j^{\rho\mu_k} K_{jk}^\rho}{\lambda D_j^{\rho\mu_k} K_{jk}^\rho + (1-\lambda) D_j^{\rho\mu_s} S_{jk}^\rho}$

$$\omega_j^{cs} = \frac{(1-c)\left(\lambda D_j^{\rho\mu_k} K_j^\rho + (1-\lambda) D_j^{\rho\mu_s} S_j^\rho\right)^{\frac{\sigma}{\rho}}}{c D_j^{\sigma\mu_u} U_j^\sigma + (1-c)\left(\lambda D_j^{\rho\mu_k} K_j^\rho + (1-\lambda) D_j^{\rho\mu_s} S_j^\rho\right)^{\frac{\sigma}{\rho}}}$$

- We make extensive use of these shares, which can be measured empirically from
  - capital share $\frac{vK_j}{Y_j} = \omega_j^c \omega_j^{cs}$
  - unskilled labor share $\frac{w_j^u U_j}{Y_j} = 1 - \omega_j^{cs}$

- Though we treat them as predetermined here, robustness checks account for the potential endogeneity of these shares

A Decomposition

- Cost minimization with respect to S and U plus full differentiation yields a relative inverse demand equation

changes in the skill bias of agglomeration economies

changes in the relative supply of skill

changes in capital intensity: >0 with K-S complementarity (endogenous)

$$d \ln \frac{w_j^s}{w_j^u} = \sigma d(\mu_s - \mu_u) \ln D_j + (\sigma - 1) d \ln \left(\frac{S_j}{U_j} \right) + (\sigma - \rho) \omega_j^c d \ln \left(\frac{K_j}{S_j} \right)$$

$$+ (\sigma - \rho) \omega_j^c d(\mu_k - \mu_s) \ln D_j + \sigma d \ln \left(\frac{A_s}{A_u} \right) + (\sigma - \rho) \omega_j^c d \ln \left(\frac{A_k}{A_s} \right)$$

Interaction between K-S complementarity and changes in factor bias of agglom economies

Factor-Biased Technical Change Terms

- Standard two-factor models with skilled and unskilled labor would have only the second term, with a coefficient of

    1/(elasticity of substitution between skilled & unskilled labor)

- This equation forms the basis for one estimating equation

Understanding Capital Intensity

$$d\ln\left(\frac{K_j}{S_j}\right) = \frac{1}{1-\rho} d\ln\left(\frac{w_j^s}{v}\right) + \frac{\rho}{1-\rho} d\left(\mu_k - \mu_s\right)\ln D_j + \frac{\rho}{1-\rho} d\ln\left(\frac{A_k}{A_s}\right)$$

- Relative price effect
- Agglomeration Effect
- Factor-biased technical change effect

Associated Reduced Form

$$d\ln\frac{K_j}{S_j} = \frac{(1-\sigma)(1-\omega_j^{cs})d(\mu_s-\mu_u) + [(1-\sigma)\omega_j^{cs}\omega_j^c + (\sigma-\rho)\omega_j^c + \rho]d(\mu_s-\mu_k) - d\mu_s}{(\sigma-\rho)\,\omega_j^c(1-\omega_j^{cs}) - (1-\rho)(1-\omega_j^c\omega_j^{cs})}\ln D_j$$
$$+\frac{(1-\sigma)(1-\omega_j^{cs})}{(\sigma-\rho)\omega_j^c(1-\omega_j^{cs}) - (1-\rho)(1-\omega_j^{cs}\omega_j^c)}[d\ln\frac{S_j}{U_j} + d\ln\left(\frac{A_s}{A_u}\right)] \tag{3}$$
$$+\frac{[d\ln v - d\ln A_j - d\ln A_s] - [(1-\sigma)\omega_j^{cs}\omega_j^c + (\sigma-\rho)\omega_j^c + \rho]d\ln(A_k/A_s)}{(\sigma-\rho)\omega_j^c(1-\omega_j^{cs}) - (1-\rho)(1-\omega_j^{cs}\omega_j^c)}$$

- Coefficients in the second and third terms are negative -> given capital-skill complementarity, skilled wages increase with
    - Declines in the price of capital (Krusell et al. story)
    - Increases in TFP or productivity of skilled labor
    - Decreases in the relative supply of skill

- Sign of first term is ambiguous, though it increases as agglomeration economies become more skill biased, raising skilled wages

- Broad lesson is that capital-skill complementarity interacts with agglomeration economies & relative labor factor supplies

Estimating Equation #2

$$d\ln\frac{K_j}{Y_j} = \frac{(\rho-\sigma)(1-\omega_j^c)(1-\omega_j^{cs})d(\mu_s-\mu_u)+(\sigma\omega_j^{cs}\omega_j^c+(\rho-\sigma)\omega_j^c-\rho)d\mu_k}{(1-\sigma)\omega_j^{cs}\omega_j^c+(\sigma-\rho)\omega_j^c+\rho-1}\ln D_j^c$$

$$+\frac{(\rho-\sigma)(1-\omega_j^c)(1-\omega_j^{cs})}{(\sigma-\rho)\omega_{j,t-1}^c+(1-\sigma)\omega_{j,t-1}^{cs}\omega_{j,t-1}^c+\rho-1}[d\ln\frac{S_j}{U_j}+d\ln\left(\frac{A_s}{A_u}\right)]$$

$$+\frac{1-\omega_j^{cs}\omega_j^c}{(\sigma-\rho)\omega_{j,t-1}^c+(1-\sigma)\omega_{j,t-1}^{cs}\omega_{j,t-1}^c+\rho-1}[d\ln v-d\ln A_j-d\ln A_k]$$

$$+d\ln\frac{A_k}{A_s}-d\ln A_j$$

- This is directly derived from the K/S equation. We express it in this way:
  - To be able to evaluate the sign of the second coefficient directly in a linear model → it is positive only if $\sigma > \rho$, or there exists capital-skill complementarity
  - So that the outcome can be measured exclusively with manufacturing census data with a timing that matches up

- The coefficient on $\ln D_j$ is 0 if the change in agglomeration economies is factor unbiased or $\sigma = \rho$

Estimating Equation #3

- Cost minimization with respect to K and S yields:

$$d\ln w_j^s = \rho d\left(\mu_s - \mu_k\right)\ln D_j + (1-\rho)\, d\ln\left(\frac{K_j}{S_j}\right) + d\ln v - \rho d\ln\left(\frac{A_k}{A_s}\right)$$

- Holding dln(K/S) fixed, larger cities will have greater increases in the skilled wage if agglomeration economies become more skill biased relative to capital biased, regulated by the relevant elasticity of substitution

- dln(K/S) is substituted with the equation given above

Empirical Implementation of Labor Supply Conditions

- From our earlier treatment, we have the following supply equation for each skill group to each CBSA, which can be thought of as a "reduced form" of the structural relative supply equation above

$$\Delta_t \ln N_j^g = \delta_t + \alpha_1 \Delta_t \ln \widehat{N}_j^g + \alpha_2 \ln N_{jt-1}^{g,imm} + \alpha_3 \ln D_j + u_{jt}$$

- Here, in principle one should control for all exogenous variables that influence labor wages and local prices
  - If "instrument" is a good one, it will not be related to these things, so leaving out some such controls (like local housing supply elasticity) will not affect the results

- Calculate components of the instrument as:

$$\hat{N}_{jt}{}^{g} = \sum_c \frac{N_{j70}{}^{gc}}{N_{70}{}^{gc}} N_t{}^{gc} \text{ so } \Delta_t \ln \hat{N}_j{}^{g} \approx \sum_c \frac{N_{j70}{}^{gc}}{N_{j70}{}^{g}} d \ln N^{gc}$$

- Following Lewis (2011), also control for the lagged relative quantities of skilled versus unskilled immigrants in all equations
  - This control never influences the results in application described below

“Supply” Estimates by Education
Manufacturing Workers

	Δln(Quantity of Workers With Indicated Education)				
	< HS	HS	Some Coll.	College	>College
Δln(Predicted	0.38***	0.23***	0.080	0.15***	-0.040
Quantity)	(0.044)	(0.037)	(0.050)	(0.055)	(0.097)
ln(CBSA Population)	-0.11***	-0.041***	-0.032**	-0.075***	-0.0038
	(0.012)	(0.016)	(0.013)	(0.016)	(0.022)
ln(Immigrants of	0.059***	-0.0043	-0.015	0.034***	0.0070
indicated educ.$_{t-1}$)	(0.0078)	(0.012)	(0.0092)	(0.010)	(0.015)
Observations	2,752	2,765	2,707	2,374	2,424
R-Squared	0.25	0.22	0.66	0.48	0.21
Year FE	Yes	Yes	Yes	Yes	Yes

- “First stage” identification will come mostly from variation in U

Resulting Fourth Estimation Equation

- The additional estimation equation to the three demand side equations, which can be thought of as a relative supply equation or a "first stage" equation

$$\Delta_t \ln \frac{S_j}{U_j} = \delta_t + \alpha_1 \Delta_t \ln \frac{\widehat{S}_j}{\widehat{U}_j} + \alpha_2 \ln \frac{S_{jt-1}^{imm}}{U_{jt-1}^{imm}} + u_{jt}$$

- Here α_1 is our key first stage coefficient of interest and has a statistically significant value of 0.21 for raw units and 0.17 for efficiency units

Identification

- Idea is to exploit the decline in S/U in Los Angeles during the 1990s both because a lot of Mexican immigrants came during that decade and because there were a lot of Mexicans in Los Angeles in 1970 -> Need to condition flexibly on time effects

- That is, for this empirical strategy to be valid, it must be true in the context of the model that

    - The relative stocks of immigrants from each country in 1970 is not correlated with changes in CBSA TFP in the 1980s, 1990s or since 2000 conditional on city size

- In a more general sense

    - There cannot be unobservables correlated with 1970 immigrant shares by education across CBSAs that predict changes in wages or capital share, conditional on city size

IV Results

	Raw Counts			
	Δln(S/U) F.S.	Δln(K/Y)	Δln(w^s)	Δln(w^s/w^u)
Δln(Predicted S /Predicted U)	0.21***			
	(0.045)			
Δln(Skilled Labor /Unskilled Labor)		0.64***	-0.25**	-0.43***
		(0.20)	(0.11)	(0.099)
ln(CBSA Population)	0.0012	0.0076	0.0094***	0.013***
	(0.0065)	(0.0078)	(0.0026)	(0.0020)
ln(Skilled Imm. / Unskilled Imm.)$_{t-1}$	0.044***	-0.033***	0.021**	0.026***
	(0.012)	(0.013)	(0.0100)	(0.0088)
Year = 2000	-0.31***	-0.19*	0.12**	0.24***
	(0.017)	(0.11)	(0.060)	(0.059)
Year = 2005-2007	-0.45***	0.059	0.12***	0.081***
	(0.030)	(0.049)	(0.025)	(0.022)
Constant	0.51***	-0.19***	-0.027***	0.064***
	(0.026)	(0.014)	(0.0077)	(0.0067)
Observations	2,751	2,047	2,751	2,751
First stage F		20.0	21.7	21.7

Information about substitutability of factors

Evidence of capital-skill complementarity, as in Lewis (2011)

Evidence of increase in the skill bias of agglomeration economies

Relationship to Existing Evidence

- Interpreting these wage gap results in the context of a two factor model implies an elasticity of substitution of 1.6 to 4.2

    - It has been growing over time. Estimates by decade yield 1.4 for the 1980s, 2.6 for the 1990s and 3.6 since 2000. Therefore, our estimates are in line with those in the literature

- Our evidence echoes that in the literature in support of the existence of capital-skill complementarity

- Ours is the first direct evidence of the increasing skill bias of agglomeration economies

Limitations of Linear IV

- While it is possible to recover the existence of relevant forces at play, reduced form linear equations do not deliver any parameter estimates

- The theory tells us that the coefficients are heterogeneous in nonlinear functions of shares ω^{c} and ω^{cs}, meaning that even with the exogenous variation available in the skill ratio, some particular unidentifiable local average treatment effects are being identified

- Useful to separate out previous coefficients into structural components

“Structural” Estimation

- Treat the 3 structural equations plus the first-stage equation literally
- Augment all structural equations with
  - The additional linear control for lagged immigrant skill mix in the “first stage” equation
  - Time fixed effects fully interacted with ln D_j in each eqn (“sparse” model)
    - Does not allow us to recover estimates of the agglomeration parameters but does allow us to recover estimates of the elasticities of substitution
  - Time fixed effects only in each equation to account for differences in the change in the capital rental rate & average TFP in different time periods (“full” model)
    - Allows for recovery of all non-share parameters
- Estimate using feasible generalized nonlinear least squares
  - IV still solves endogeneity problem b/c eqn is linear in the endogenous variable

Identification

- Compare two cities of the same sizes that had different exogenous relative labor supply shocks
    - This comparison identifies elasticity of substitution parameters
- Compare two cities of different sizes that received the same exogenous relative labor supply shock
    - This comparison identifies parameters that govern changes in the factor bias of agglomeration economies

Structural Estimates

Parameter	Description	Sparse Model Counts	Sparse Model Eff Unit	Full Model Counts	Full Model Eff Unit
α_1	Coefficient on instrument in Equation (6)	0.28	0.24	0.24	0.20
		(0.05)	(0.05)	(0.04)	(0.04)
σ	$1/(1-\sigma)$=elast of sub btw K or S and U	0.85	0.90	0.84	0.87
		(0.03)	(0.02)	(0.02)	(0.02)
ρ	$1/(1-\rho)$=elast of sub btw K and S	0.22	0.43	-0.51	-0.61
		(0.26)	(0.23)	(0.42)	(0.62)
$d\mu_s$	change in skilled labor biased agglom.			0.017	0.017
				(0.002)	(0.002)
$d\mu_k$	change in capital biased agglom.			-0.014	-0.011
				(0.003)	(0.002)
$d\mu_u$	change in unskilled labor biased agglom.			-0.009	-0.004
				(0.003)	(0.002)
$d\ln(A_s/A_u)$	skill biased technical change			0.217	0.183
				(0.065)	(0.066)

- Evidence of capital-skill complementarity ($\sigma > \rho$)
    - increasing skill-bias of agglomeration economies
    - declining capital bias of agglomeration economies
- Unreported time fixed effects indicate thatdlnv-E[dlnA] < 0 for all 3 study periods
- Skill biased technical change

A Decomposition

- Revisit this equation

$$
\begin{aligned}
d\ln\frac{w_j^s}{w_j^u} \;=\; & \sigma d(\mu_s-\mu_u)\ln D_j+(\sigma-1)d\ln\left(\frac{S_j}{U_j}\right)+(\sigma-\rho)\omega_j^c d\ln\left(\frac{K_j}{S_j}\right) \\
& +(\sigma-\rho)\omega_j^c d(\mu_k-\mu_s)\ln D_j+\sigma d\ln\left(\frac{A_s}{A_u}\right)+(\sigma-\rho)\omega_j^c d\ln\left(\frac{A_k}{A_s}\right)
\end{aligned}
$$

- Note that in bringing this equation to the data, we have augmented it such that we did not estimate it directly, but can still recover its components

- Our analysis is better suited to understanding how important these mechanisms are for understanding the increasingly positive relationship between the log skill gap and city size than for understanding changes in wage gaps overall

    - The augmented components that we have added for identification purposes (year FE and immigrant skill ratios) are important for understanding national trends in wage gaps

Decomposition Results

- Regress skill gap components on time fixed effects andln(D)

- All coefficients are significant

Actual		0.015	0.012
Predicted		0.016	0.013
Predicted Using K,S,U from Data		0.015	0.014
Agglomeration	Eqn 2, Term 1	0.021	0.018
S-U Shifts	Eqn 2, Term 2	0.001	0.000
K-S Shifts	Eqn 2, Term 3	0.011	0.014
Agglom. K-S Complem. Interaction	Eqn 2, Term 4	-0.018	-0.018
Skill-Biased Technical Change	Eqn 2, Term 5	0.000	0.000
Capital-Biased Technical Change	Eqn 2, Term 6	-0.001	0.000

- Model fits this element of patterns in the data well

- Discrepancy between the two Predicted versions are the use of exogenous vs. all variation in dln(S/U)

- Clear central role for the increased skill bias of agglomeration economies

Why Have Cities Become More Capital Intensive?

- Decomposition of dln(K/S)
- These results are shown using a calibrated value of r because it is estimated imprecisely

Calibrated value of ρ			0.24	0.28
3	Agglomeration	Eqn 3, Term 1	-0.003	0.000
4	Agglom, S Bias	Eqn 3, Term 1, $d\mu_s \neq 0$	0.015	0.015
5	Agglom, K Bias	Eqn 3, Term 1, $d\mu_k \neq 0$	-0.017	-0.015
6	Agglom, U Bias	Eqn 3, Term 1, $d\mu_u \neq 0$	-0.001	0.000
7	S-U Shifts	Eqn 3, Term 2 (1st pt)	0.007	0.006
8	Skill-Biased Technical Change	Eqn 3, Term 2 (2nd pt)	0.004	0.003
9	Prod. + Capital Price	Eqn 3, Term 3	-0.002	-0.002

From Wages to Income, Well-Being and Neighborhoods

- Moretti (2013) argues that trends in price differences across cities means that from a measurement perspective, growth in real wage inequality has been slower than growth in nominal wage inequality

- What about the microstructure of metropolitan areas

  - Chetty et al. (2014, 2015) provide evidence that neighborhoods and youth environment matter

  - There is a lot we don’t know about why and how they matter

  - Neighborhood dynamics can be important for understanding trends and persistence in inequality

Patterns of Neighborhood Dynamics (Baum-Snow & Hartley, 2015)

		Inequality Criterion		
		Fraction	Fraction	Mean HH
Period		White	College Ed	Income
1970-1980	Constant	1.001	1.114	0.883
		(0.014)	(0.008)	(0.017)
1980-1990	ΔLn(Employment),	-0.080	-0.036	-0.115
	standard devs.	(0.032)	(0.013)	(0.055)
	Constant	0.976	1.109	0.910
		(0.011)	(0.005)	(0.025)
1990-2000	Constant	0.934	1.056	0.896
		(0.012)	(0.006)	(0.006)
2000-2010	ΔLn(Employment),	-0.043	-0.009	-0.082
	standard devs.	(0.023)	(0.011)	(0.025)
	Constant	0.869	1.002	1.006
		(0.011)	(0.005)	(0.009)
0	ΔLn(Employment),	-0.123	-0.085	-0.155
	standard devs.	(0.053)	(0.026)	(0.064)
	Constant	0.773	1.184	0.846
		(0.021)	(0.012)	(0.026)

Conclusions and Research Ideas

- Agglomeration economies exist, but what are the main mechanisms driving them?
    - Potential "static" forces: better market access for intermediate goods, lower cost inputs due to economies of scale, fewer frictions to technology adoption, ???
    - Potential "dynamic" forces: more rapid human capital accumulation (but how?), differently operating internal labor markets; these forces seem to be worker and not firm specific
- Local labor markets matter for changes in wage inequality
    - What mechanisms are driving more rapid increases in the price of skill in larger cities?
        - Rise of the service sector interacted with agglomeration economies in service provision?
        - Differences in skill-complementary capital costs?
- What about the microstructure of metropolitan areas and neighborhood dynamics?