Global Value of Cities

Aakash Bhalothia, Gavin Engelstad, Gaurav Khanna, Harrison Mitchell

The Effects of Cities

- "*Mostly in the world there aren't poor people. There are people in poor places*" (Pritchett, 2009)

The Effects of Cities

- "*Mostly in the world there aren't poor people. There are people in poor places*" (Pritchett, 2009)

- Question 1: What is the Causal Effect of Each City in the World?
  - What are the earning gains in moving from Paris to LA? Bangalore to San Francisco?
  - Place effects vs sorting?
    - City Effects (productivity, types of jobs, technology, infrastructure, etc.)
    - Sorting on individual abilities (Young 2013)

The Effects of Cities

- "*Mostly in the world there aren't poor people. There are people in poor places*" (Pritchett, 2009)

- Question 1: What is the Causal Effect of Each City in the World?
  - What are the earning gains in moving from Paris to LA? Bangalore to San Francisco?
  - Place effects vs sorting?
    - City Effects (productivity, types of jobs, technology, infrastructure, etc.)
    - Sorting on individual abilities (Young 2013)

- Question 2: Why are Some Cities More Productive?
  - Correlates of city premia?

The Effects of Cities

- "*Mostly in the world there aren't poor people. There are people in poor places*" (Pritchett, 2009)

- Question 1: What is the Causal Effect of Each City in the World?
  - What are the earning gains in moving from Paris to LA? Bangalore to San Francisco?
  - Place effects vs sorting?
    - City Effects (productivity, types of jobs, technology, infrastructure, etc.)
    - Sorting on individual abilities (Young 2013)

- Question 2: Why are Some Cities More Productive?
  - Correlates of city premia?
  - Size of potential gains from migrating from low to high-wage cities?
  - Do these migration gains vary through the economic development process?

Challenges and Solutions

1. Data: comparable panel data from different countries, with history of moves
   - Global coverage of high-skilled workers (LinkedIn)
   - Work histories of 513mn people in 220K cities

2. Endogenous sorting across places
   - Movers design to identify causal effect of places
   - Tests for pre-trends, symmetry, and firm heterogeneity

3. Interpreting city premia, implications
   - Compare premiums within and across countries
   - Correlates of city effects: population, industrial structure?
   - How distribution of city effects (within country) vary with development

Contribution

1. Causal effects of places (Card et al., 2024; Combes et al., 2008; de la Roca and Puga, 2017)
    - Global analysis
    - Applied econometric questions related to estimating place effects
    - Event study designs

Contribution

1. Causal effects of places (Card et al., 2024; Combes et al., 2008; de la Roca and Puga, 2017)
   - Global analysis
   - Applied econometric questions related to estimating place effects
   - Event study designs

2. Gains from Migration (Pritchett and Hani, 2020)
   - Comparing both; movement across cities instead of just countries
   - Why do gains vary over the development process?

Contribution

1. Causal effects of places (Card et al., 2024; Combes et al., 2008; de la Roca and Puga, 2017)
   - Global analysis
   - Applied econometric questions related to estimating place effects
   - Event study designs

2. Gains from Migration (Pritchett and Hani, 2020)
   - Comparing both; movement across cities instead of just countries
   - Why do gains vary over the development process?

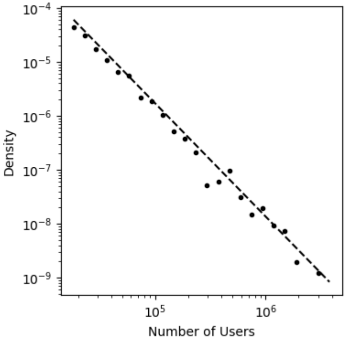
3. Allocation across economic units explain cross-country differences in TFP (Hsieh and Klenow, 2009 for Firms; Gollin et al., 2014 for Sectors)
   - Why are some cities more productive than others?
   - Aggregate gains from re-allocation across cities and firms

Data

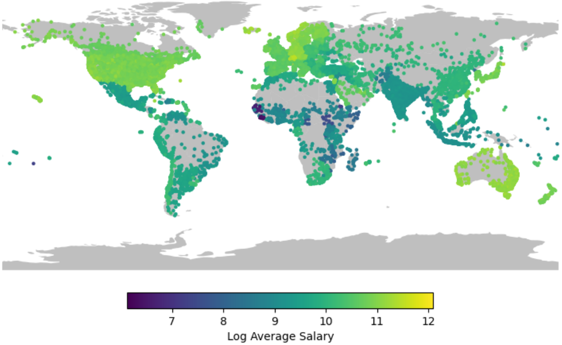
Employment histories of over 513 million individuals from 220,000 cities in 191 countries (LinkedIn)

- Education (degree, institution, location, time)
- Raw Job title: mapped to occupations (8-digit O*NET), Seniority
- Start Date, End Date
- Company
- Raw location - matched to city, country
- Salary (Imputed)
- ▸ Sample Linkedin Profiles ▸ Salaries by Occupation ▸ Salaries by Seniority

Coverage: Map of wages of cities $>$ 100 users



(a) Cities by Number of Users



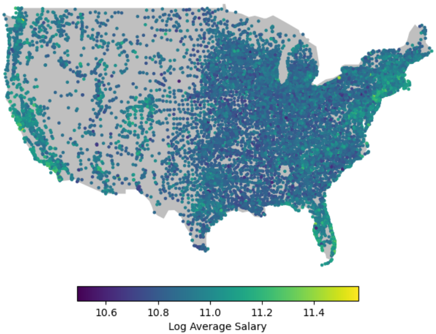
(b) Average Log City Wages, Global

Imputed Salary Variable

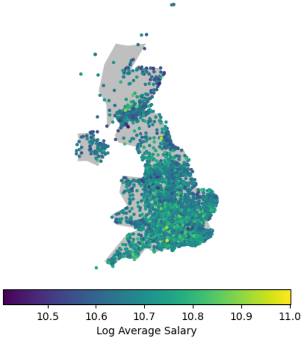
- Function of company, occupation title, tenure, year, and city characteristics (median housing value, population density, unemployment rate, poor share)
- Relevant welfare measure - high score indicates 'better' jobs
  - Amanzadeh et al. (2024) validate with individual Glassdoor data, and national surveys
  - Highly correlated with GDP per capita ▸ Graph
- Under standard movers-design assumptions: Valid estimated place effects on imputed salary
- We argue imputed wages can lead to biased city effects
  - Develop new methods to bound estimates when data are imputed
  - Quantify the degree of bias

Average Salaries by City

(a) USA

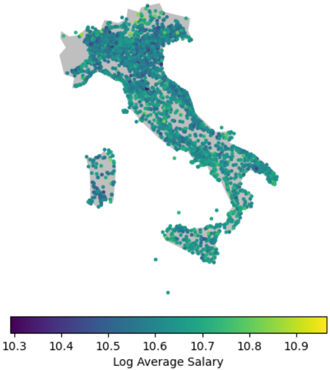


(b) UK

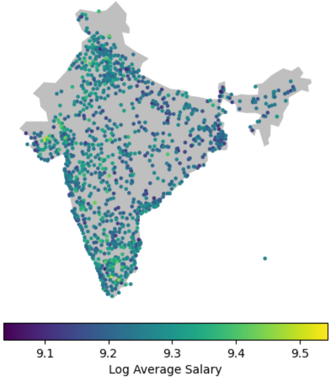


Average Salaries by City

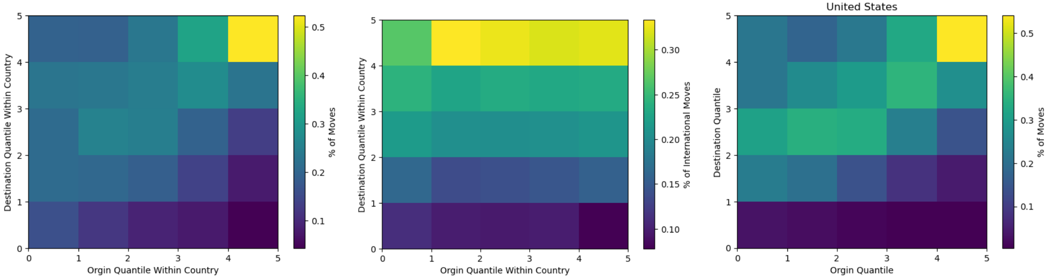
(a) Italy



(b) India



Transition Matrices



(a) Within Country Moves (b) Cross Border Moves (c) US: Transition Matrix

Event Study Analysis (Finkelstein et al 2016)

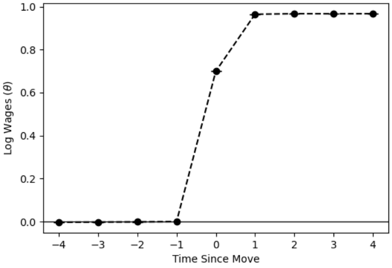
$$y_{it} = \alpha_i + \tau_t + I_{r(i,t)} + \theta_{\mathrm{r}(i,t)} \delta_i I_{r(i,t)} + \eta_{it}$$

- $\delta_i = \bar{y}_{d(i)} - \bar{y}_{o(i)}$
- $I_{r(i,t)}$ Vector of relative years. α_i individual fixed effects. τ_t year fixed effects.
- Plot $\theta_{\mathrm{r}(i,t)}$ – relative year coefficients.

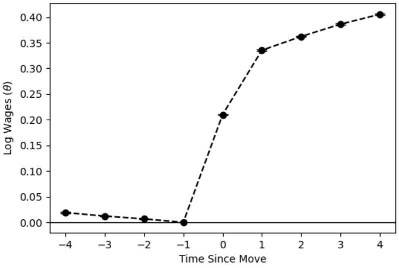
Event Study Analysis (Finkelstein et al 2016)

$$y_{it} = \alpha_i + \tau_t + I_{r(i,t)} + \theta_{\mathrm{r}(i,t)}\delta_i I_{r(i,t)} + \eta_{it}$$

- $\delta_i = \bar{y}_{d(i)} - \bar{y}_{o(i)}$
- $I_{r(i,t)}$ Vector of relative years. α_i individual fixed effects. τ_t year fixed effects.
- Plot $\theta_{\mathrm{r}(i,t)}$ – relative year coefficients.

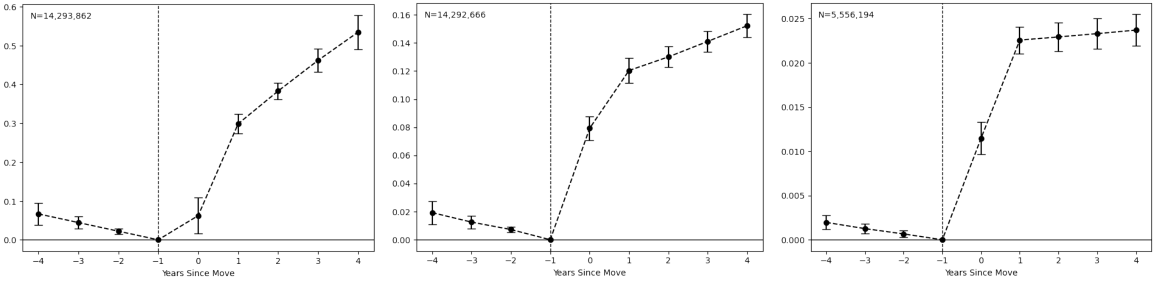


(a) International Movers, N = 2,736,283



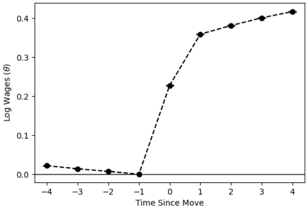
(b) Internal Movers, N = 13,222,406

Seniority, Occupation, and Industries

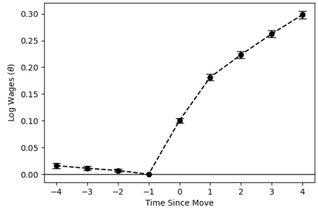


(a) Seniority (b) Occupation Score (c) Industry Score

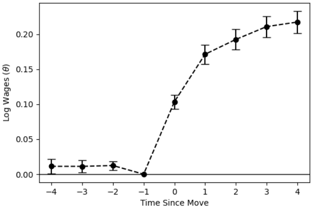
Salary Effects for Internal Migration



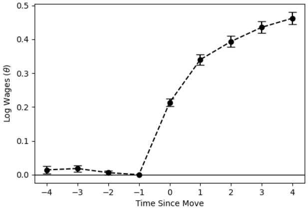
(a) USA, N = 5,869,926



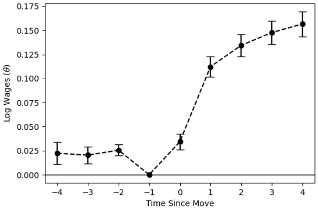
(b) UK, N = 832,197



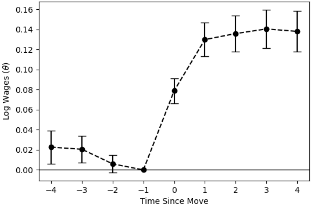
(c) Italy, N = 332,031



(d) Germany, N = 433,050



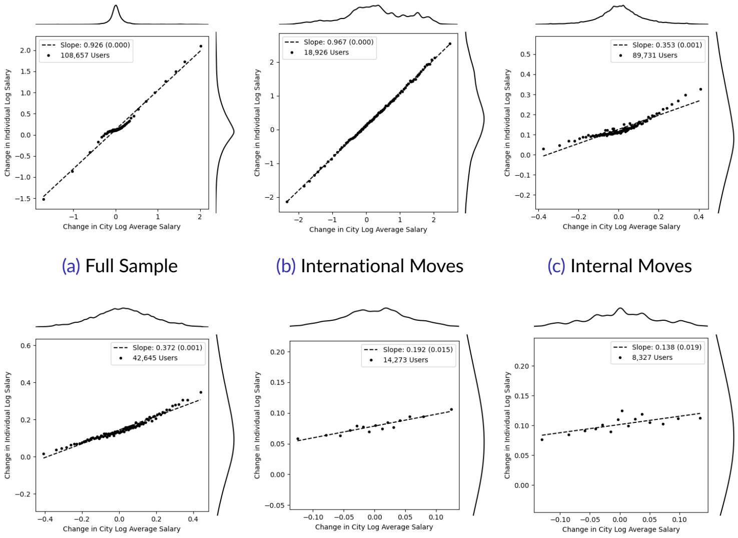
(e) India, N = 774,142



(f) Mexico, N = 202,888

Gain by Pairwise City Wages

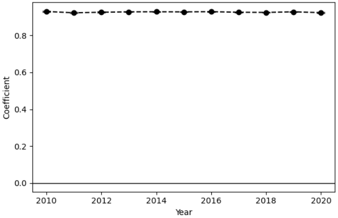
Gain by Pairwise City Wages



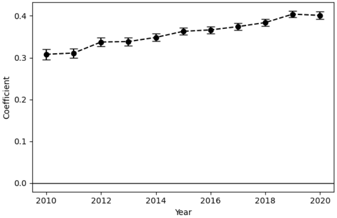
(a) Full Sample (b) International Moves (c) Internal Moves

(d) USA (e) Italy (f) Mexico

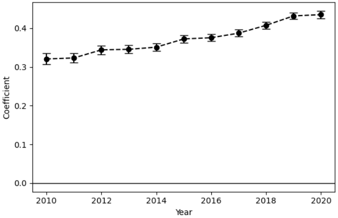
Changes in Gains Over Times



(a) International Moves



(b) Internal Moves



(c) USA Moves

Basic AKM

- Decompose to worker effects (α_i), city effects ($\psi_{\boldsymbol{J}(i,t)}$), and observables

$$Log(Earnings)_{it} = \alpha_i + \tau_t + \psi_{\boldsymbol{J}(i,t)} + x'_{it}\beta + \epsilon_{it}$$

- What is allowed: Systematic mobility on individual and city characteristics
  - e.g., Productive workers are mobile
  - e.g., Productive cities are in demand (or trending; when relative-year FEs)
  - e.g., Assortative matching: productive workers move to productive cities

Basic AKM

- Decompose to worker effects (α_i), city effects ($\psi_{\boldsymbol{J}(i,t)}$), and observables

$$Log(Earnings)_{it} = \alpha_i + \tau_t + \psi_{\boldsymbol{J}(i,t)} + x_{it}'\beta + \epsilon_{it}$$

- What is allowed: Systematic mobility on individual and city characteristics
    - e.g., Productive workers are mobile
    - e.g., Productive cities are in demand (or trending; when relative-year FEs)
    - e.g., Assortative matching: productive workers move to productive cities
- Assumptions:
    - Log additive separability of worker and city effects
        - Multiplicative in levels
    - Exogenous mobility
        - No sorting on individual-city match quality
        - Drift in individual effects (shocks or human capital growth) or transitory factors.
        - Job loss; bid up wages from outside or time-to-adaptation (before move), one-time signing bonus

Testing Assumptions and Hierarchy Effects

- Event Studies:
  - Pre-trends
  - No spikes before / after moves (signing bonus)
  - α_i allow for non-movers to be different.
  - Individual time-since-move fixed effects allow trends around moves

- Symmetry in moves between pairs of cities
  - Additive Separability
  - Sorting on Match Quality

Testing Assumptions and Hierarchy Effects

- Event Studies:
  - Pre-trends
  - No spikes before / after moves (signing bonus)
  - α_i allow for non-movers to be different.
  - Individual time-since-move fixed effects allow trends around moves
- Symmetry in moves between pairs of cities
  - Additive Separability
  - Sorting on Match Quality
- Perhaps still concerned about moving on match quality...
- Hierarchy Effects:
  - Moving from low-wage city (high-paying firm) to a high-wage city (but low-paying firm)
  - Muddies interpretation: no longer move to a random firm

Testing Assumptions and Hierarchy Effects

- Event Studies:
  - Pre-trends
  - No spikes before / after moves (signing bonus)
  - α_i allow for non-movers to be different.
  - Individual time-since-move fixed effects allow trends around moves
- Symmetry in moves between pairs of cities
  - Additive Separability
  - Sorting on Match Quality
- Perhaps still concerned about moving on match quality...
- Hierarchy Effects:
  - Moving from low-wage city (high-paying firm) to a high-wage city (but low-paying firm)
  - Muddies interpretation: no longer move to a random firm
  - Card et al. (2024): Estimate firm effects $\gamma_{\boldsymbol{f}(i,t)}$, and aggregate to city effects Γ_j:

$$y_{it} = \alpha_i + \tau_t + \gamma_{\boldsymbol{f}(i,t)} + x'_{it}\beta + \epsilon_{it}$$

$$\Gamma_j = \frac{\sum_{j(f)=j} N_f \gamma_f}{\sum_{j(f)=j} N_f}$$

Expected Bias from Predicted Wages

- Wages are imputed: can cause bias if assortative matching
- Amanzadeh et al. (2024) validate with individual-level Glassdoor data ▸ Graph
- What bounds can we put on wage effects?

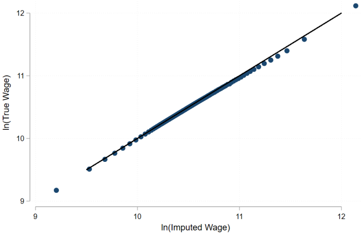
Expected Bias from Predicted Wages

- Wages are imputed: can cause bias if assortative matching
- Amanzadeh et al. (2024) validate with individual-level Glassdoor data ▸ Graph
- What bounds can we put on wage effects?

- Methods we develop are useful when using imputed wages (Abramitzky et al., 2012; 2021)
- High wage imputation even in gold-standard data like ACS and CPS (30% – Bollinger & Hirsch 2013)

Expected Bias from Predicted Wages

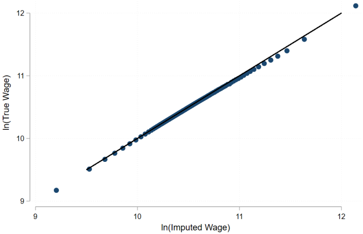
- Use employer-employee data from Italy to compute imputed wages with similar information (e.g., firm, tenure, salary code)
- Imputed and real wages highly correlated ($R^2 = 0.875$)



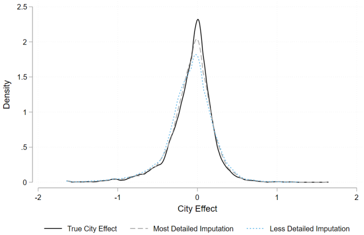
(a) Binned scatter of true vs. imputed wages (b) City effects under different imputations (c) Excess variance under added noise

Expected Bias from Predicted Wages

- Use employer-employee data from Italy to compute imputed wages with similar information (e.g., firm, tenure, salary code)
- Imputed and real wages highly correlated ($R^2 = 0.875$)
- Imputation overstates wage variation explained by cities by 12%



(a) Binned scatter of true vs. imputed wages

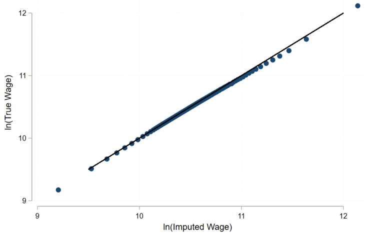


(b) City effects under different imputations

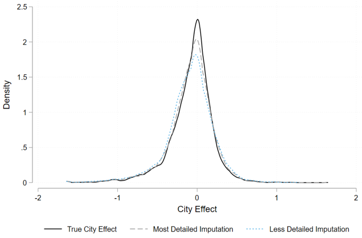
(c) Excess variance under added noise

Expected Bias from Predicted Wages

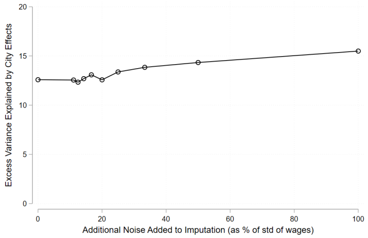
- Use employer-employee data from Italy to compute imputed wages with similar information (e.g., firm, tenure, salary code)
- Imputed and real wages highly correlated ($R^2 = 0.875$)
- Imputation overstates wage variation explained by cities by 12%
- What if we added more and more noise? (e.g., developing countries worse data)



(a) Binned scatter of true vs. imputed wages



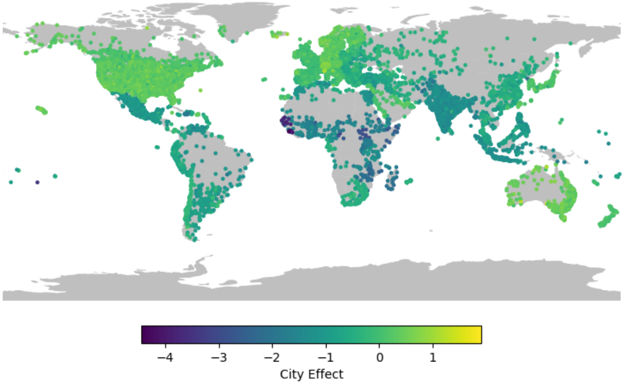
(b) City effects under different imputations



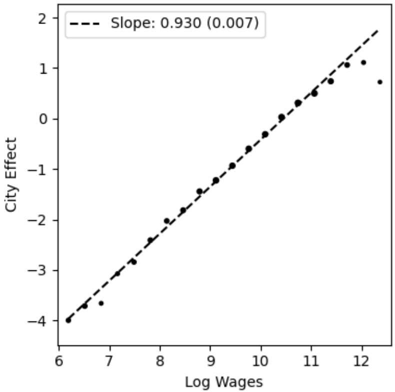
(c) Excess variance under added noise

City Effects: Global

City Effects: Global

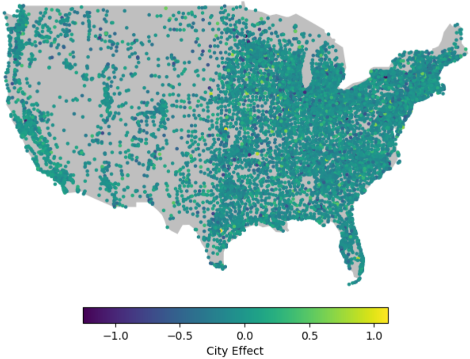


(a) Estimated City Effect

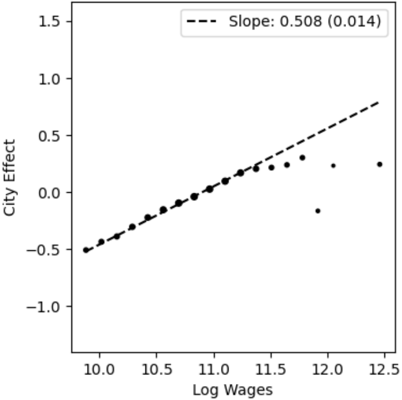


(b) Log City Wages vs City Effect

City Effects: USA

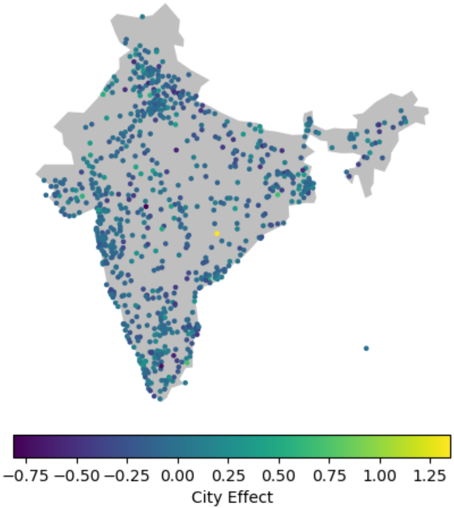


(a) Estimated City Effect

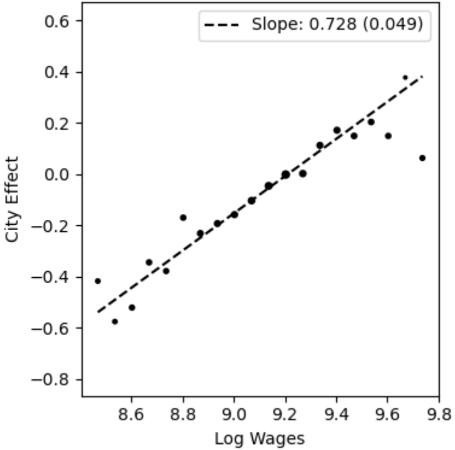


(b) Log City Wages vs City Effect

City Effects: India



(a) Estimated City Effect



(b) Log City Wages vs City Effect

Contribution of City Effects

$$\Omega \equiv \frac{Cov(\bar{y}_j, \Gamma_j)}{Var(\bar{y}_j)} \text{ Share of Variance in Wages Due to City Effects}$$

$$S(R, R') = \frac{\Gamma_R - \Gamma_{R'}}{\bar{y}_R - \bar{y}_{R'}} \text{ The share of wage differences explained by city effects}$$

Contribution of City Effects

$$\Omega \equiv \frac{Cov(\bar{y}_j, \Gamma_j)}{Var(\bar{y}_j)}$$ Share of Variance in Wages Due to City Effects

$$S(R, R') = \frac{\Gamma_R - \Gamma_{R'}}{\bar{y}_R - \bar{y}_{R'}}$$ The share of wage differences explained by city effects

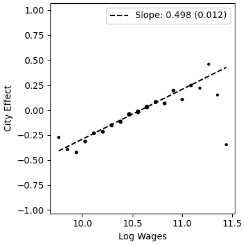
	Below to Above Median (1)	Bottom to Top 25% (2)	Bottom to Top 5% (3)	Bangalore to San Francisco (4)	Var $(\bar{y}_j)$ (5)
Difference in Ln Wages	0.75	1.27	2.39	1.1	–
Difference due to City	0.65	1.07	2.05	1.02	–
Share due to City	0.87	0.84	0.86	0.93	0.93
Bounded Share					(0.75)

Decomposition of Wage Differences

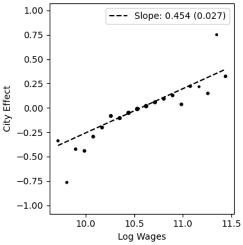
Panel A: United States	Below to Above Median (1)	Bottom to Top 25% (2)	Bottom to Top 5% (3)	San Diego to New York (4)	Var $(\bar{y}_j)$ (5)
Difference in Ln Wages	0.32	0.52	1.0	0.2	–
Difference due to City	0.14	0.24	0.45	0.04	–
Share due to City	0.45	0.45	0.45	0.22	0.51
Bounded Share					(0.41)

Panel B: India	Below to Above Median (1)	Bottom to Top 25% (2)	Bottom to Top 5% (3)	Kolkata to Bangalore (4)	Var $(\bar{y}_j)$ (5)
Difference in Ln Wages	0.26	0.44	0.84	0.13	–
Difference due to City	0.17	0.29	0.47	0.1	–
Share due to City	0.65	0.65	0.56	0.74	0.73
Bounded Share					(0.59)

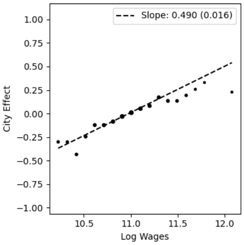
City Effects, Other Countries



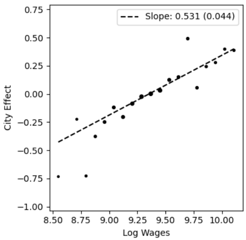
(a) UK



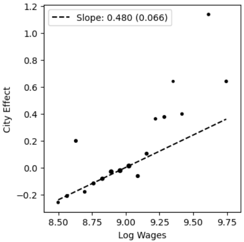
(b) Italy



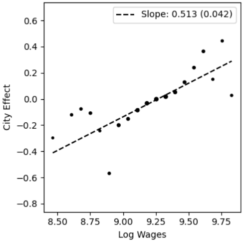
(c) Germany



(d) Mexico



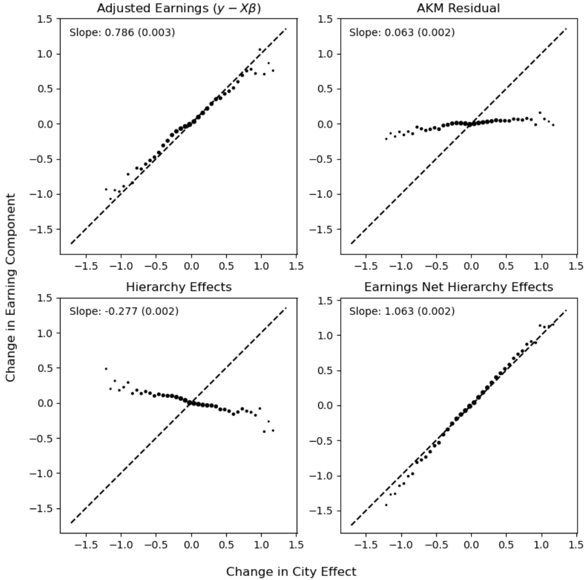
(e) Nigeria



(f) Philippines

Hierarchy Effects

Hierarchy Effects



Explaining City Effects: The Allocation of Workers

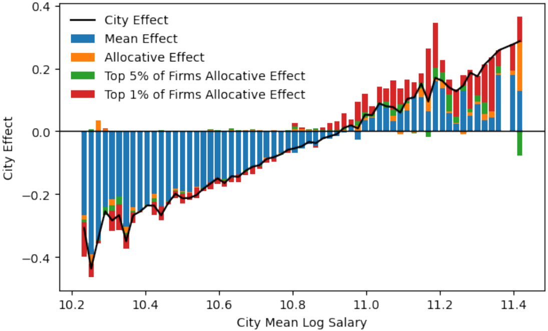
Decompose city effects: how much is due to the employment allocation across high and low-productive firms (Olley & Pakes 1996):

$$\Gamma_j = \sum_{j(f)=j} s_f \gamma_f = \bar{\gamma}_j + \sum_{j(f)=j} (s_f - \bar{s}_j)(\gamma_f - \bar{\gamma}_j)$$

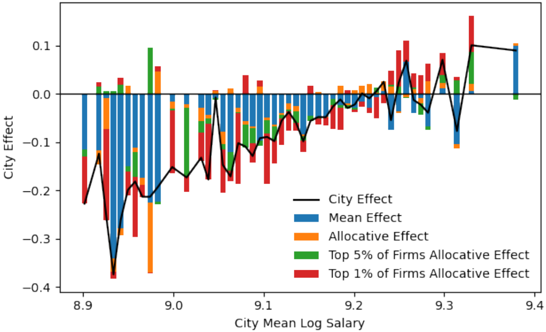
Explaining City Effects: The Allocation of Workers

Decompose city effects: how much is due to the employment allocation across high and low-productive firms (Olley & Pakes 1996):

$$\Gamma_j = \sum_{j(f)=j} s_f \gamma_f = \bar{\gamma}_j + \sum_{j(f)=j} (s_f - \bar{s}_j)(\gamma_f - \bar{\gamma}_j)$$



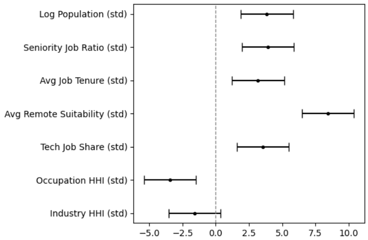
(a) USA



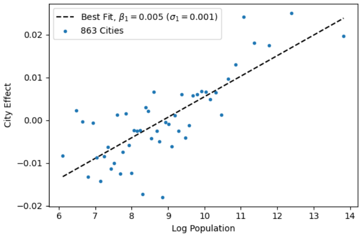
(b) India

What Explains Higher City Effects?

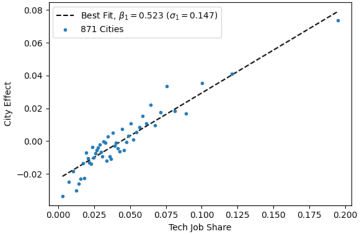
What Explains Higher City Effects?



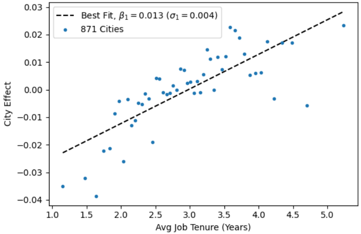
(a) Correlates of City Effects



(b) Population



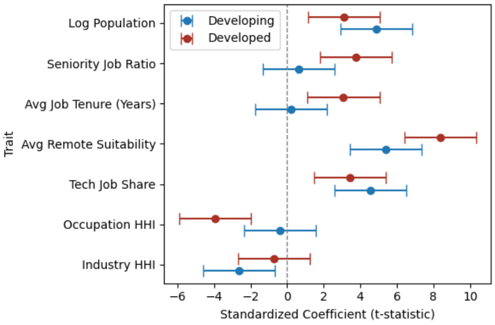
(c) Tech Job Share



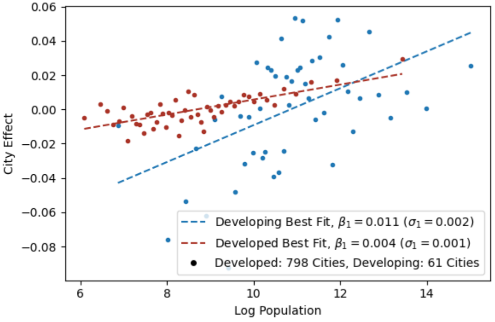
(d) Average Job Tenure

City Effect Correlates by Development Status

City Effect Correlates by Development Status

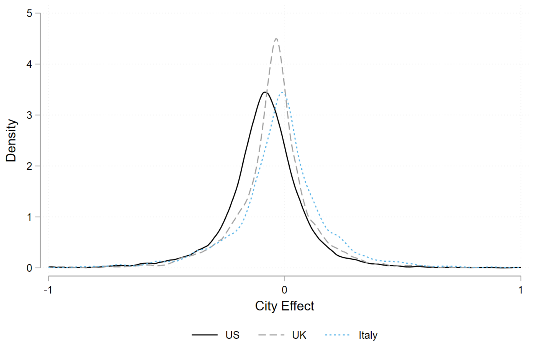


(a) By Development Status

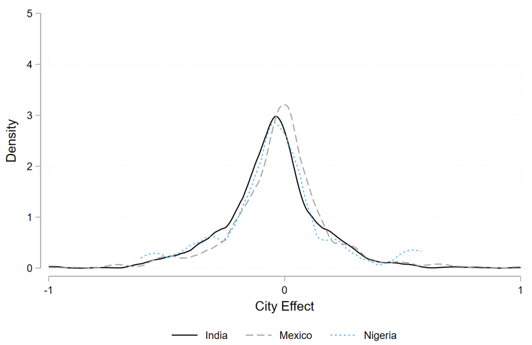


(b) Population

Distribution of City Effects



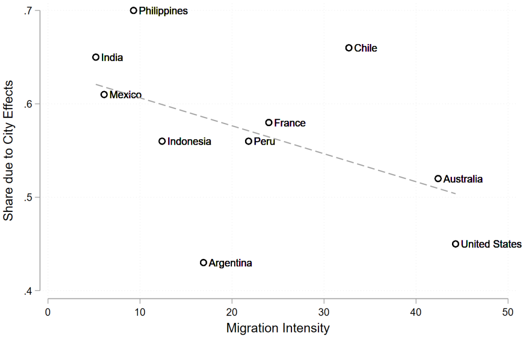
(a) City Effects Distribution, Developed Countries



(b) City Effects Distribution, LMICs

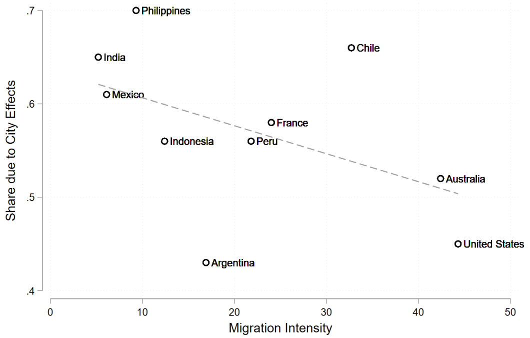
Share Due to City Effects And Migration Frictions

Share Due to City Effects And Migration Frictions

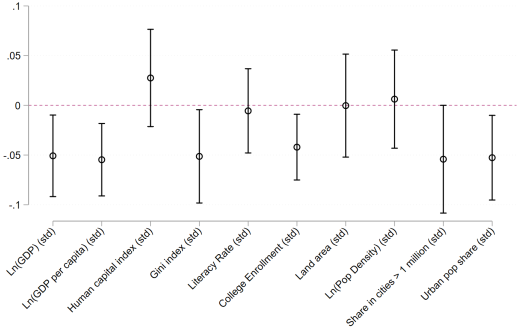


(a) Migration Intensity (Bell et al. 2015) and City Effects

Share Due to City Effects And Migration Frictions



(a) Migration Intensity (Bell et al. 2015) and City Effects



(b) Country Aggregates and Median Share

Takeaways for high-skilled migration

- **Internal Migration:** Reallocation of workers can increase average incomes
  - Variation across countries in potential gains (across development status)
- **Cross-border Migration:** Difference in levels across borders is so large
  - ‘Moving someone from *any city* in India to Houston would have similarly large gains in earnings’

Takeaways for high-skilled migration

- **Internal Migration:** Reallocation of workers can increase average incomes
  - Variation across countries in potential gains (across development status)
- **Cross-border Migration:** Difference in levels across borders is so large
  - 'Moving someone from *any city* in India to Houston would have similarly large gains in earnings'
- **City Effects Correlated With** Agglomeration and Skills
- **Next steps?** What data would be useful to release?

Thank you

Example LinkedIn Profile

Experience

Tobin Pre-Doctoral Fellow
Yale University · Full-time
Aug 2020 - Jun 2021 · 11 mos
New Haven, Connecticut, United States

Research fellow under Professor Rohini Pande

Research Associate
NERA Economic Consulting
Jul 2019 - Jul 2020 · 1 yr 1 mo
Greater New York City Area

Research Assistant
Department of Economics, UC Berkeley
May 2018 - May 2019 · 1 yr 1 mo
Berkeley, California

Evaluated the effects of unconditional direct cash transfers using data from a randomized control trial in Kenya (working with Professor Edward Miguel)

Undergraduate Student Instructor
University of California, Berkeley
Jan 2018 - May 2019 · 1 yr 5 mos

Teaching Assistant for an intermediate Data Science course (Data Science 100)

Software Engineering Intern
NVIDIA
May 2017 - Aug 2017 · 4 mos
Santa Clara, CA

Built tools to uncover performance bottlenecks in CUDA compiler and applications

Back

Example of a Bad LinkedIn Profile

Experience

Research Assistant
Yale University
Jun 2019 - Present · 5 yrs
New Haven, Connecticut

Supplemental Instructor
Macalester College
Aug 2018 - May 2019 · 10 mos
Greater Minneapolis-St. Paul Area

Farm Hand
Easy Bean Farm
May 2018 - Sep 2018 · 5 mos
Greater Minneapolis-St. Paul Area

Macalester College
2 yrs 10 mos
Greater Minneapolis-St. Paul Area

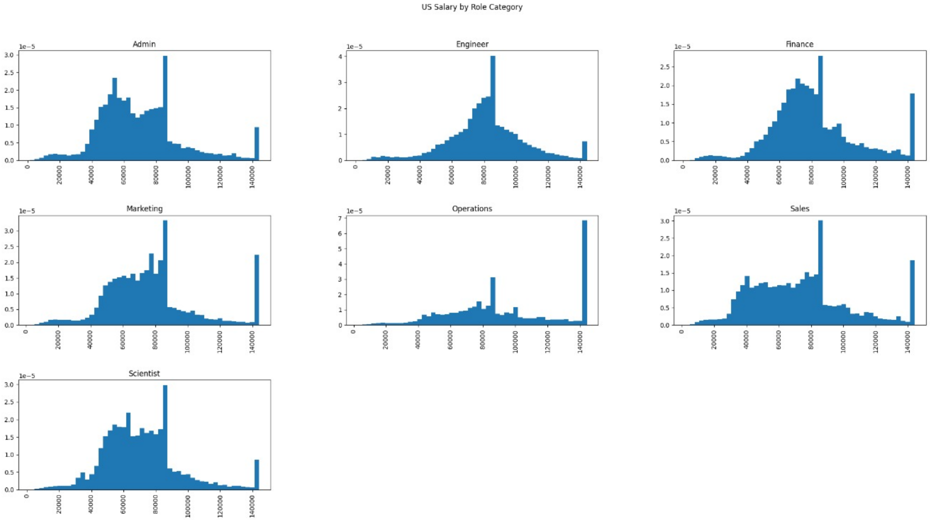
Preceptor
Jan 2018 - May 2018 · 5 mos

Studeny Recyler
Aug 2015 - May 2018 · 2 yrs 10 mos

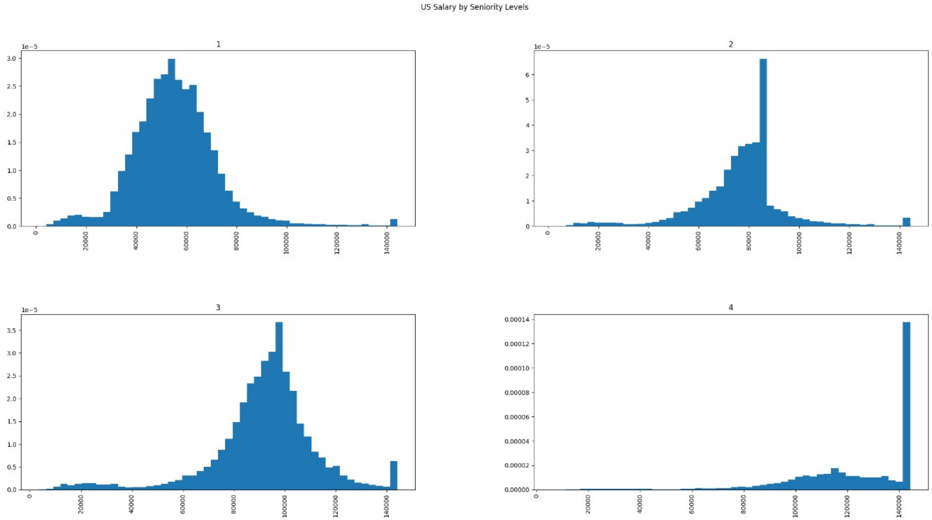
- Responsible for maintaining clean living area in dorms
- Ensures classrooms are orderly before each academic day
- Practices sustainable cleaning methods

Back

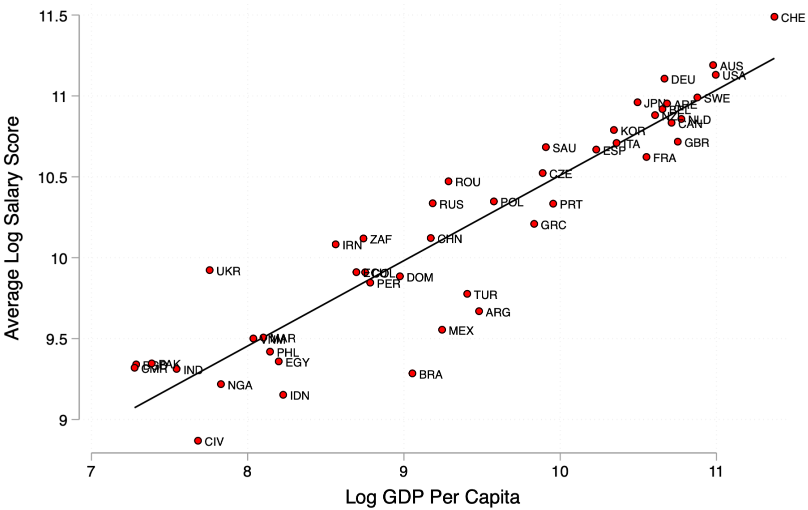
US Salaries by Occupation



US Salaries by Seniority



Salary Score and Per Capita GDP



Back

Going from Salary Score to Salary

- With true salaries

$$Log(Earnings)_{it} = \alpha_i + \Phi_{C(i,t)} + X_t + \epsilon_{it}$$

- With salary score

$$P_{it} = \hat{\alpha}_i + \hat{\Phi}_{C(i,t)} + \hat{X}_t + \hat{\epsilon}_{it}$$

- Need: $\forall C, \hat{\Phi}_{C(i,t)} = \Phi_{C(i,t)}$

- Surrogacy Assumption:

$$Earnings_{it} \perp\!\!\!\perp \Phi_{C(i,t)} \mid P_{it}, \alpha_i, X_t$$

  - Violation: Cities have an effect on salaries even after controlling for firm, occupation title, tenure, city characteristics, individual characteristics, and time-fixed effects

- Under the above assumption, ATE on P_{it} = ATE on $Earnings_{it}$ (Athey et al., 2024; Prentice, 1989)

- Validate with US data

(Back)