The Geography of Automation

Md. Deen Islam

Department of Economics

University of Dhaka

December 26, 2022

Introduction

- Automation and AI are replacing human labor for a wide range of routine-type tasks.
    - Serious consequences for employment and wages of occupations performing these tasks.
- This chapter studies the effects of automation on **regional employment and wage-inequality**.
- The aim of chapter 1 is:
    - To understand the impact of automation on the geographic distribution of employment.
    - To investigate the effects of automation on regional wage inequality.

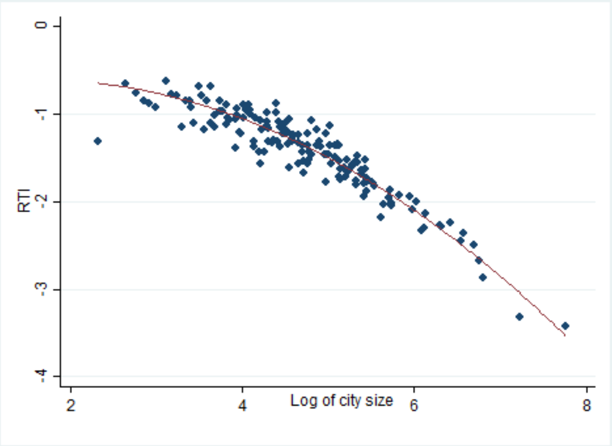
The Geography of Automation

- Automation can lower labor requirements and increase firms' incentives to locate in larger cities
  - Large cities offer productivity advantages, but wages in large cities are high (Melo et al. 2009).
  - Labor intensive firms locate in low-wage small cities (Broersma and Oosterhaven 2009, Farzmand, Tu, and Norcio 2015).
- Automation can alter the distribution of employment and wages across regions.

Contribution to Literature

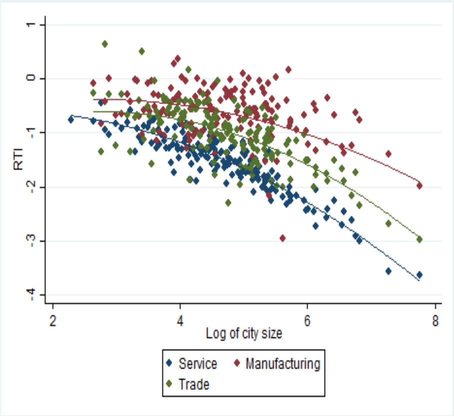
- Measuring inequality: Diamond (2016), Ganong and Shoag (2017), Giannone (2017), Harrigan et al. (2018), Cerina et al. (2019), Eckert (2019), Moll, Rachel and Restrepo (2021).
    - Develop a spatial general equilibrium model to measure regional wage disparity.
- Agglomeration of firms: Ciccone and Hall (1996), Ellison and Glaeser (1999), Combes et al. (2010), Gaubert (2018). Davis and Dingel (2019, 2020), Davis et al.(2021).
    - Identify automation as one of the determinants of agglomeration of firms in large urban areas.

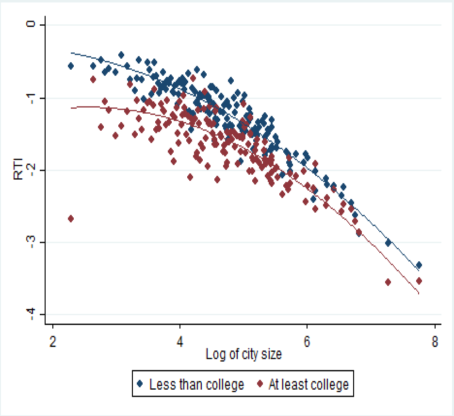
Fact 1: Workers in larger cities perform less routine-type task



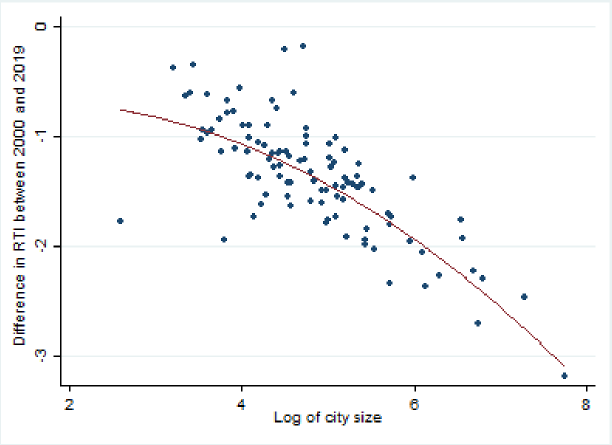
Data Panel regression: Fact 1

Fact 1 holds for different levels of education and for different sectors



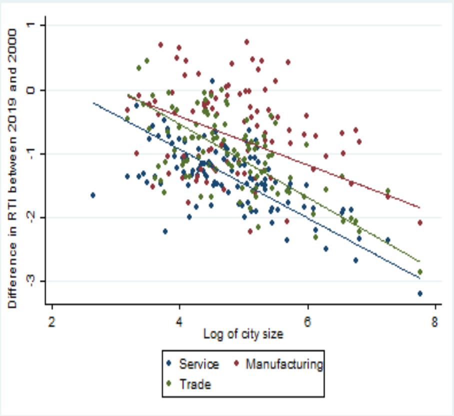


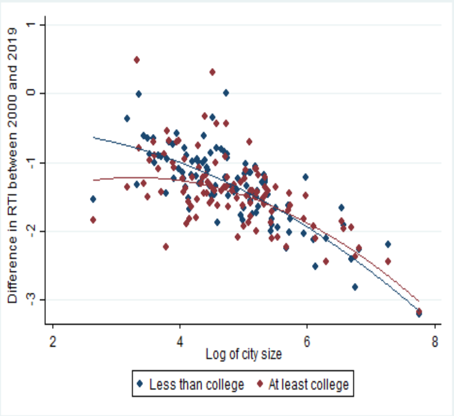
Fact 2: Larger cities experienced greater fall in routine-type task.



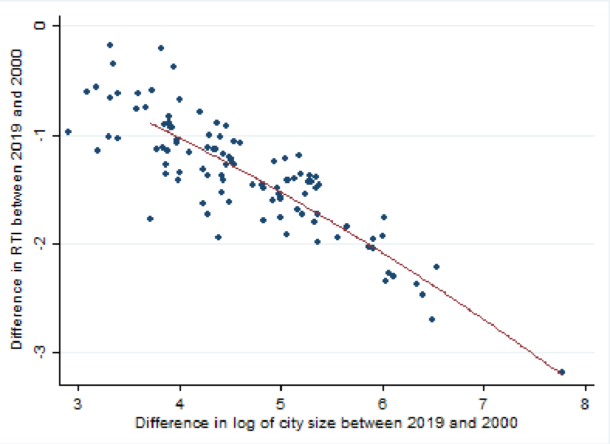
Panel regression: Fact 2

Fact 2 holds for different levels of education and for different sectors.



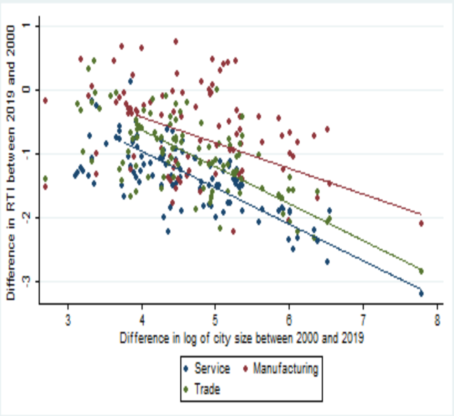


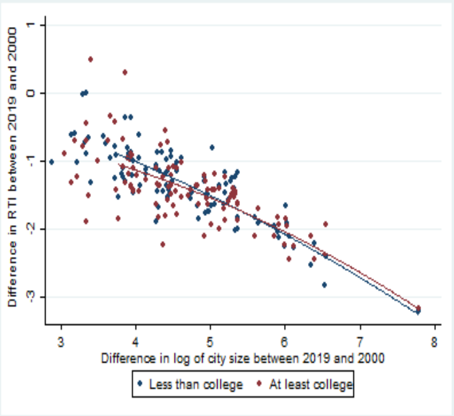
Fact 3: Faster growing cities experienced rapid fall in routine-type task.



Regression: Fact 3

Fact 3 also holds for different levels of education and for different sectors.





Stylized facts: Summary

- Firms in large urban areas use less routine-type task.
- Routine-type task falls at a faster rate in larger cities.
    - Re-organization of economic activities over space.
    - Increase in regional wage inequality.
- To investigate these implications in a robust way, I develop a spatial model in the next section.

The model

- The model combines agents sorting into different locations as in Gaubert (2018) and Desmet et al. (2018) with the task-based production technology similar to Acemoglu and Restrepo (2018).
- More automated firms use less routine-type tasks and sort into larger cities.
- Greater level of automation for all firms lead to regional wage inequality.

Model Environment

- An economy with a total population L distributed over N cities indexed by n: $\sum_{n=1}^{N} L_n = L$.
- Each agent supplies 1 unit of labor inelastically and can freely move across cities.
- A city n has an endogenously determined housing supply H_n.

Consumer's problem

- Agent i living in city n has preferences over housing (H_i) and a tradable composite good (X_i) characterized by the following utility function:

$$u_n^i(H,X) = a_n^i H_{ni}^{\beta} X_{ni}^{(1-\beta)}$$

where

- β= Expenditure share on housing, ($\beta \in (0,1)$)
- a_n^i= idiosyncratic preference of agent i for city n: $Pr\left(a_n^i \leq a\right) = e^{-a^{-\nu}}$.
- X_{ni}= CES aggregate of varieties: $X_{ni} = \left[\int_{\omega\in\Omega} x_i(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right]^{\frac{\sigma-1}{\sigma}}$
  - Ω=Exogenous mass of varieties,
  - σ=Elasticity of substitution between varieties. ($\sigma > 1$)

Consumer's location choice

- The log of the Indirect Utility function is:

$$v_n^i = w_n - \beta r + ln(a_n^i)$$

- Agent's optimal city choice is the solution of:

$$\max_n \quad v_n^i = w_n - \beta r + ln(a_n^i)$$

- The labor supply of the city n is then

$$L_n = \frac{W_n^\nu}{\sum_{n=1}^N W_n^\nu} L$$

  - W_n = Wage in city n
  - R = Rent in city n

Production technology

- Each variety ω is produced by a unique firm by combining a unit continuum of tasks.

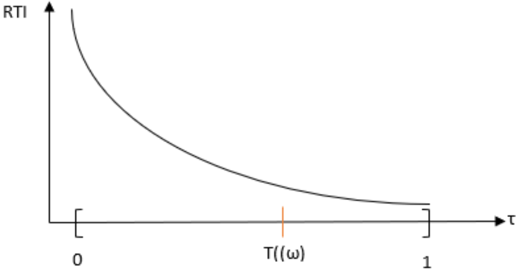
- The production function for ω in n is:

$$q_n(\omega) = Z_n^{\omega} \min_{\tau} \Big\{ y(\omega, \tau) \Big\}$$

  - $Z_n^{\omega} = z_n^{\omega} A_n L_n^{\theta}$,
    - z_n^{ω}=City-specific exogenous firm productivity,
    - A_n=Exogenous city productivity,
    - L_n=City size.
  - $y(\omega, \tau)$= Quantity of the task τ in ω

Tasks and automation

- Tasks are ranked by $\tau \in [0, 1]$, where $\tau = 0$ for most routine-type task, and $\tau = 1$ for least routine-type task.
- RTI index and task index τ in ω is related as:



- $T(\omega)$ is the automation frontier up to which all routine-type tasks can be automated.

Production technology for tasks

- Tasks indexed lower than $T(\omega)$ can be produced using either labor or capital.
- Tasks indexed higher than $T(\omega)$ can be produced only using labor.
- Production function for a task τ in ω is:

$$y(\omega,\tau) = \begin{cases} \gamma l(\tau) + \eta k(\tau), & \tau \leq T(\omega) \\ \gamma l(\tau), & \tau > T(\omega) \end{cases}$$

  - η=Capital productivity,
  - γ=Labor productivity.

Firm's location choice

- A firm chooses a production location that minimizes its unit cost:

$$n^*(T(\omega)) = \underset{n}{argmin}\left\{ \frac{1}{Z_n^\omega}\left(\frac{T(\omega)P_k}{\eta} + \frac{(1-T(\omega))W_n}{\gamma} \right) \right\}$$

  - P_k=Price of capital,

- The total labor demand in city n:

$$l_n^d = \int_\omega \frac{(1-T(\omega))q(\omega)}{\gamma Z_n^\omega} d\omega, \qquad \omega \in \Omega_n$$

- The mass of firms in city n:

$$\Omega_n = \left\{ \omega : \frac{(1-T(\omega))W_n}{\gamma Z_n^\omega} = min\left\{ \frac{(1-T(\omega))W_n}{\gamma Z_{n'}^\omega} \right\}, n' = 1, 2, \ldots, N \right\}$$

Capital market

- One unit of capital is produced by using ι quantity of output.
- Price of capital is : $P_k = \iota \times P$.
- Aggregate capital demand is:

$$K = \sum_n \int_\omega \frac{T(\omega)q(\omega)}{\eta Z_n^\omega} d\omega$$

Housing market

- Housing supply is assumed to be perfectly elastic.
- The housing market equilibrium condition:

$$L_n \frac{\beta W_n}{R} = H_n$$

Spatial Equilibrium

- **Definition:** The equilibrium in this spatial economy is a menu of city level wages (W_n), allocation of labor L_n and distribution of firms across cities as functions of A_n, z_n^{ω}, $T(\omega)$ and ι that satisfies the following conditions:

    1. Individuals choose a city to reside in to maximize the indirect utility.
    2. Firms choose a production location to minimize their unit costs of production.
    3. Local labor market clears.
    4. Capital is optimally allocated.
    5. The housing market clears.

Automation and City Size

- **Proposition 2:** *In cross-section, L_n and $T_n = E\{T(\omega)|A_n\}$ are increasing in A_n. Therefore, larger cities have lower RTI.*

- Cities with larger A_n have lower unit cost of production, especially for firms with high $T(\omega)$.

Proposition 2 & 3

Automation and Inequality

- **Proposition 3:** *A uniform increase in the level of the automation potentials for all firms, $T'(\omega) > T(\omega)\ \forall\ \omega$,*
  - *(i) Increases the number of firms in cities with larger A_n.*
  - *(ii) Increases the T_n in cities with larger A_n.*
  - *(iii) Increases wage dispersion across cities.*
- Intuition:
  - A uniform increase in $T(\omega)$ leads to greater sorting of firms into cities with larger A_n increasing wage and number of firms in those cities.

Proposition 2 & 3

Model summary

- The model provides predictions about the relationship between A_n, L_n, and $T(\omega)$
- It predicts the effects of automation on city size, and wage dispersion across cities.
- I estimate the model quantitatively to measure the magnitude of these effects.

Parameters of the model

- To estimate the model quantitatively, I fix the parameter values:
  - Calibration: $\beta, \theta, \sigma, \gamma, \eta$.
  - Estimation: $\nu, \iota, b, std(A_n), std(z_n^\omega)$.

Calibration of parameter values

Name	Definition	Value	Source
β	Expenditure share on housing	0.35	consumer expenditure–2019, 2019 BLS
θ	Agglomeration elasticity	0.1	Redding and Turner (2015)
σ	Elasticity of substitution	3.0	Anderson and van Wincoop (2004), Giovanni and Levchenko (2012)
γ	Labor productivity	0.75	Brynjolfsson and Hitt (2003)
η	Capital productivity	0.22	Brynjolfsson and Hitt (2003)

Estimation of parameter values

Name	Definition	Value	Data source
ν	Shape parameter of Frećhet	1.5	IPUMS, AHS, and US census
ι	Capital production parameter	1.44	Compustat
b	$T(\omega) \sim U(0, b)$	0.6	IPUMS
$std(A_n)$	Std. Dev. of A_n	0.28	Compustat and IPU
$std(z_n^\omega)$	Std. Dev. of z_n^ω	0.02	Compustat

ν ι b $std(A_n)$ $std(z_n^\omega)$ Result

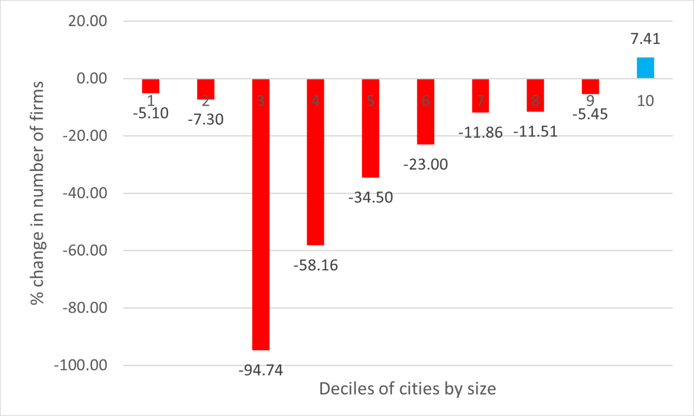
Counterfactual analysis

- The baseline and counterfactual distributions of the automation frontiers are:

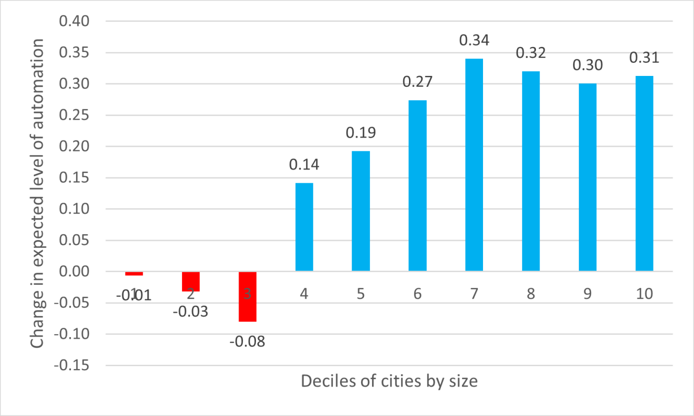
  1. Baseline automation frontier: T $\sim$ U(a, b) $\sim$ U(0, 0.6),
  2. Counterfactual automation frontier: T1 $\sim$ U(0, 0.7),

- From 2000 to 2019, *b* has increased by about 0.08, so an increase in *b* by 0.1 indicates a change of about 25 years.

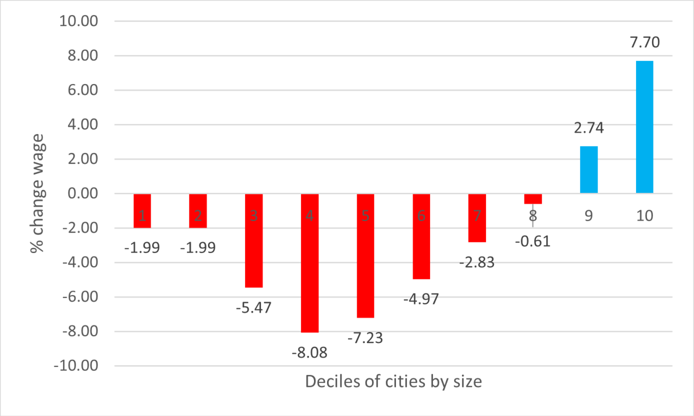
Cities in the largest deciles gains whereas cities in all other deciles loss number of firms



Larger cities have positive change in the level of Expected automation



Wage dispersion rises following an uniform increase in automation potential



Proposition 2 & 3

Concluding remark

- This paper
  - Provided new stylized facts on the relationship between city size and automation
  - Developed a general equilibrium spatial model to investigate the effects of automation on the distribution of economic activity and regional inequality.
  - Equilibrium analysis of the model indicated that an increase in automation leads to wider dispersion in wages and welfare across cities.

Stylized facts: Data sources

- Individuals' occupation, wage, industry, location and other characteristics from IPUMS USA
- Score for routine, abstract, and manual tasks from Autor and Dorn (2013)
- Industry and occupation crosswalks between different censuses from Dorn (2009)

Summary statistics

Variables

- Measure of automation: Routine Task Index in city n (RTI_n):
  - $RTI_n = lnRS_n - lnAS_n - lnMS_n$
  - $RS_n = \frac{\sum_i routine\_score_{i,n}}{\sum_i Emp_{i,n}}$
  - $AS_n = \frac{\sum_i abstract\_score_{i,n}}{\sum_i Emp_{i,n}}$
  - $MS_n = \frac{\sum_i manual\_score_{i,n}}{\sum_i Emp_{i,n}}$
- City size (CS): Sum of employed individual i in city n
  - $CS = \sum_i Emp_{i,n}$
- Other variables: Median Age (MA), Median year of education (ME), Fraction of white employee (WF), Hourly wage rate (WR).

Fact 1

Summary statistics

Table: Descriptive statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
RTI	Overall	-1.616	0.617	-4.421	0.988	N = 5425
	Between		0.528	-4.045	-0.602	n = 408
	Within		0.281	-2.651	0.338	$\bar{T}$= 13.297
City size	Overall	137.878	221.026	1	3184	N = 5425
	Between		186.353	11.545	2558	n = 408
	Within		44.633	-503.872	763.878	$\bar{T}$= 13.297
Median age	Overall	41.211	3.62	25	67	N = 5425
	Between		2.312	29.5	52.4	n = 408
	Within		3.013	26.078	66.666	$\bar{T}$= 13.297
Median education	Overall	11 693	1.579	6	16	N = 5425
	Between		1.221	9.636	15	n = 408
	Within		1.045	3.46	16.755	$\bar{T}$= 13.297
White employee	Overall	0.416	0.088	0	1	N = 5425
	Between		0.075	0.079	0.543	n = 408
	Within		0.041	-0.028	0.934	$\bar{T}$= 13.297
Hourly wage	Overall	6.505	8.703	0	154.48	N = 5425
	Between		6.568	0	78.963	n = 408
	Within		4.452	-28.409	82.023	$\bar{T}$= 13.297

Data

Fact 1 regression

Table: Effect of city size on automation

	Dep. Var.= *RTI*		
	(1)	(2)	(3)
lnCS	-0.422^{***}	-0.421^{***}	-0.421^{***}
	(0.021)	(0.021)	(0.022)
lnMA		-0.096	-0.02
		(0.065)	(0.066)
lnME		-0.523^{***}	-0.429^{***}
		(0.064)	(0.066)
WF		-0.321^{**}	-0.444^{***}
		(0.157)	(0.158)
lnWR		-0.048^{***}	-0.049^{***}
		(0.009)	(0.009)
Constant	0.254^{***}	2.15^{***}	1.793^{***}
	(0.093)	(0.318)	(0.32)
Time FE	No	No	Yes
Overall R^2	0.7426	0.7535	0.7544
Observations	5425	5425	5425
Groups	408	408	408

[1] Note: Standard errors are in parentheses. Standard errors are clustered at county level.

[2] $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$.

[3] Source: Estimated using IPUMS and Autor and Dorn (2013) data.

Fact 1

Fact 2 regression

Table: Larger cities in the year 2019 experienced greater automation

	Dep. Var.= ΔRTI			
	Unweighted		*Weighted*	
	(1)	(2)	(3)	(4)
*lnCS*2019	−0.115	00.016	−0.098***	−0.048***
	(0.044)	(0.059)	(0.002)	(0.004)
*lnMA*2019		−0.35		0.665***
		(0.507)		(0.044)
*lnME*2019		−0.512*		−1.113***
		(0.270)		(0.023)
*WF*2019		1.161***		1.012***
		(0.326)		(0.025)
*lnWR*2019		−0.121**		−0.044***
		(0.059)		(0.005)
Constant	0.235	2.042	0.148***	0.016
	(0.208)	(1.995)	(0.015)	(0.188)
Overall R^2	0.0051	0.1493	0.0674	0.2472
Observations	138	138	23107	23107

[1] Robust standard errors are in parentheses. Employment in 2019 is used as weight for estimates in column (3) and (4).

[2] ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

[3] Source: Estimated using IPUMS and Autor and Dorn (2013) data.

Fact 2

Fact 3 regression

Table: Faster growing cities become more automated

	Dep. Var.= ΔRTI			
	(1)	(2)	(3)	(4)
$\Delta lnCS$	-0.468^{***}	-0.434^{***}	-0.546^{***}	-0.607^{***}
	(0.081)	(0.082)	(0.114)	(0.0124)
$\Delta lnMA$		-0.409	0.185	0.337
		(0.417)	(0.491)	(0.539)
$\Delta lnME$		-0.975^{***}	-1.024^{***}	-1.165^{***}
		(0.268)	(0.247)	(0.284)
$\Delta lnWF$		-0.005	0.003	0.002
		(0.015)	(0.014)	(0.016)
$\Delta lnWR$		-0.094^{**}	-0.129^{**}	-0.118^{**}
		(0.04)	(0.05)	(0.053)
Constant	-0.207^{***}	0.158	-0.224	-0.180
	(0.032)	(0.604)	(0.0575)	(0.668)
Overall R^2	0.2497	0.3574	0.4414	0.4603
Observations	138	138	138	138

[1] Note: Robust standard errors are in parentheses. Columns (1) and (2) show unweighted regression estimates. Column (3) provides the weighted estimates, where the county-level employment in 2000 is used as weight. Similarly, column (4) show the weighted estimates, where the county-level employment in 2019 is used as weight.

[2] $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$.

[3] Source: Estimated using IPUMS and Autor and Dorn (2013) data.

Fact 3

Estimation of ν

Table: Estimation of idiosyncratic city preferences, ν

	Dep. Var.= City level population, L_n			
	(1)	(2)	(3)	(4)
Rent	-0.601^{***}	-0.613^{***}	-0.520^{***}	-0.437^{***}
	(0.084)	(0.080)	(0.081)	(0.133)
WR		0.128^{***}	0.121^{***}	0.108^{***}
		(0.000)	(0.016)	(0.014)
MedAge			-0.01^{**}	-0.008^{**}
			(0.004)	(0.003)
MedEduc			-0.036^{***}	-0.029^{*}
			(0.008)	(0.007)
WhiteFrac			-0.951^{*}	-0.029^{***}
			(0.547)	(0.488)
Constant	9.15^{***}	9.07^{***}	9.68^{***}	8.999^{***}
	(0.559)	(0.530)	(0.570)	(0.869)
TimeFE	No	No	No	Yes
Overall R^2	0.095	0.006	0.061	0.095
N	3888	3888	3888	3888

- The estimates of the coefficient of rent range from -0.44 to -0.61, which implies a value of ν ranging from 1.26 to 1.75. I set the value of ν to 1.5.

Estimation of parameters

Estimation of ι

Table: Estimation of capital production technology, ι

	Dep. Var.= Total revenue (R)	
	(1)	(2)
Capital	0.921***	0.845***
	(0.002)	(0.034)
Labor		0.269***
		(0.032)
Constant	−0.048***	−0.976***
	(0.011)	(0.242)
R^2	0.7756	0.7698
N	76637	339

- The estimate of capital coefficient is 0.85 after controlling labor expenses.

$$\iota = \sqrt{\frac{1}{\alpha_1}\frac{K}{R}} = \sqrt{\frac{1}{0.85}\frac{6875.88}{3907.89}} = 1.44.$$

Estimation of b

- $T(\omega) \sim U(0, b)$, where b is estimated to match the variance of the distribution of RTI: $V(RTI) = \frac{b^2}{12}$
- Using IPUMS data, I estimate the variance of Industry-level RTI to be 0.029.
- The value of b is estimated to be about 0.59. I fix it to 0.6.

Estimation of parameters

Estimation of $std(A_n)$

- Following Diamond (2016), I specify:

$$lnW_n = \beta_0 + \beta_1 lneduc_n + \beta_2 RTI_n + A_n$$

- A_n is the residual of the above regression
- The estimated distribution of A_n has mean 1.032 and standard deviation 0.277.

Estimation of parameters

Estimation of $std(z_n^\omega)$

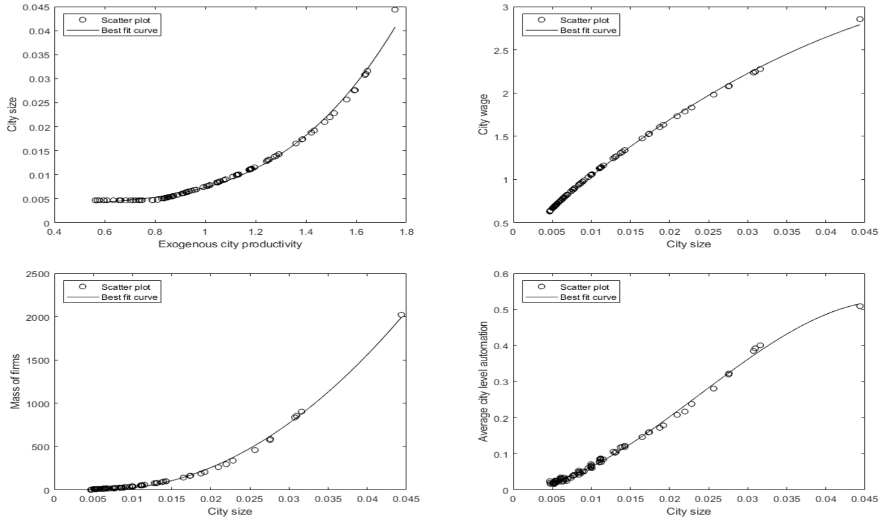
- Using the model's aggregate agglomeration benefit equation, I specify the following regression:

$$Z_n^\omega = \alpha_0 + \alpha_1 lnA_n + \theta lnL_n + z_n^\omega$$

- Estimated residuals of above regression are used as a measure of z_n^ω
- The distribution of z_n^ω has a mean of 1 and standard deviation of 0.02.

Estimation of parameters

Numerical results of the model are consistent with empirical evidence



Estimation of parameters