# Quantitative Spatial Economics III: Diamond Style Model in Urban Economics

Zibin Huang [1]

[1]College of Business, Shanghai University of Finance and Economics

April 26, 2025

# Overview


1. Introduction
2. Data and Descriptive Stats
3. Model Setting
4. Estimation
5. Solving the Equilibrium
6. Counterfactual

# Introduction

- In the next two weeks, we are going to introduce Diamond style model
- The basic idea is to transform a non-linear model to a linear one
- Then we can simply use reduced-form approach to estimate parameters one by one
- It is a good starting example for beginners to learn structural model

# Introduction

- The original model comes from Diamond (2016), who investigates how spatial sorting exacerbated inequality in the U.S.
- She finds that during the last few decades, high-skilled and low-skilled workers sorted into different locations in America

# Introduction

- High-skilled workers concentrated more and more in developed cities, increasing amenities and living costs
- Low-skilled workers were kicked out and had to live in cities with low amenities and living costs
- She shows that the overall inequality change is much larger than purely education wage gaps
- The increased inequality in amenity amplified the overall inequality

# Motivation

- China experienced high economic growth during the last forty years
- TFP and technology developed rapidly
- Did this growth lead to similar spatial sorting and gentrification?
- Now we will introduce an application of Diamond model in China's context

# Research Question

We have the following main research question:

- What is the impact of the technology shocks on migration for different skilled groups in China

We answer this question by constructing a spatial GE model with:

- Endogenous labor market
- Endogenous housing market
- Endogenous amenity supply

# Preview of Findings

- Descriptive evidence
  - Cities with stronger patent growth experienced faster wage growth for both low- and high-skilled workers
  - However, these cities attract much more low-skilled migrants than high-skilled migrants.
- Structural estimation results
  - Low-skilled workers care more about wages and housing prices, while high-skilled workers care more about amenities.
  - A positive shock in patents attracts more low-skilled workers than high-skilled workers, reduces the skill ratio and amenities, and discourages high-skilled migrants.

# Preview of Findings

- Counterfactual of reducing patent from 2015 level to 2005 level:
  - Large reduction in low-skilled migration but not high-skilled migration
  - Welfare loss for both skills, especially low-skilled people with non-ag hukou
- Generally, growth in China during 2005 to 2015 is inclusive without diversification

# Data

- Data Source:
  - Migration: micro-level Census 2005, 2010, 2015
  - Wages, housing prices, and amenities: statistical yearbooks
  - Patent citation data: China National Intellectual Property Administration and Google Patent
- Migrants: those who left hukou city for at least 6 months.
- Laborforce: age 25 - 50, currently working
- Non-Agriculture sector: working in urban regions and non-agriculture industries
- Agriculture sector: working in rural regions

# PCA Results of the Amenity Index

| | Loading | Unexplained variance |
|---|---|---|
| *Panel A: Healthcare Index* | | |
| Hospital per 10,000 residents | 0.7071 | 0.4351 |
| Doctors per 10,000 residents | 0.7071 | 0.4351 |
| *Panel B: Infrastructure Index* | | |
| Kilometers of road per 10,000 residents | 0.4178 | 0.8078 |
| Highway passengers per 10,000 residents | 0.5987 | 0.6053 |
| High-speed railway | 0.6834 | 0.4856 |
| *Panel C: Environment Index* | | |
| PM 2.5 | 0.5315 | 0.339 |
| Heavily polluted days | 0.5712 | 0.2365 |
| Polluted days | 0.6255 | 0.08434 |
| *Panel D: Education Index* | | |
| Teacher-student ratio in primary schools | 0.0818 | 0.9824 |
| Teacher-student ratio in middle schools | 0.1136 | 0.966 |
| Number of colleges | 0.5395 | 0.2333 |
| Number of Project 985 universities | 0.5886 | 0.08752 |
| Number of Project 211 universities | 0.5855 | 0.09703 |
| *Panel E: Amenity Index* | | |
| Healthcare Index | 0.6434 | 0.4391 |
| Infrastructure Index | 0.5535 | 0.5848 |
| Environment Index | -0.2340 | 0.9258 |
| Education Index | 0.4742 | 0.6952 |

# Descriptive Analysis: Spatial Distribution of Patent Shock

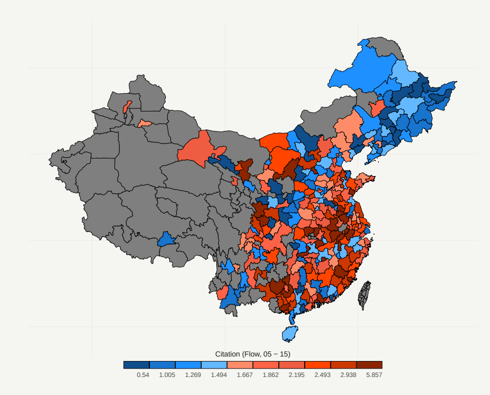



Figure 1: Spatial Distribution of $\Delta$ Log(Citation) (2005 - 2015)

Patents

# Descriptive Analysis

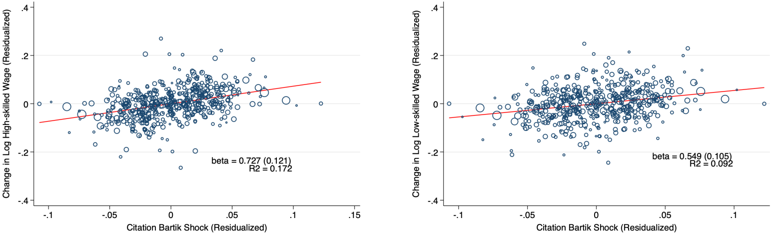



Figure 2: Effect of Citation Shock on Wages for High- and Low-skilled Workers

▸ Patents

# Descriptive Analysis

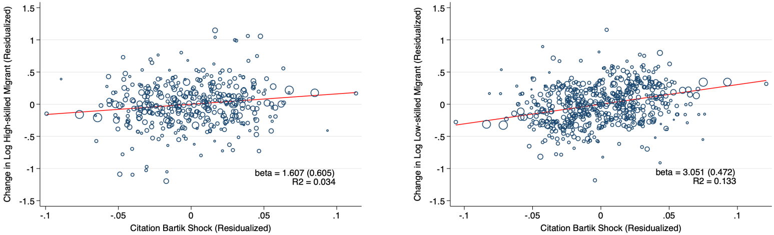



Figure 3: Effect of Citation Shock on Number of High- and Low-skilled Migrants

▸ Patents

# Descriptive Analysis

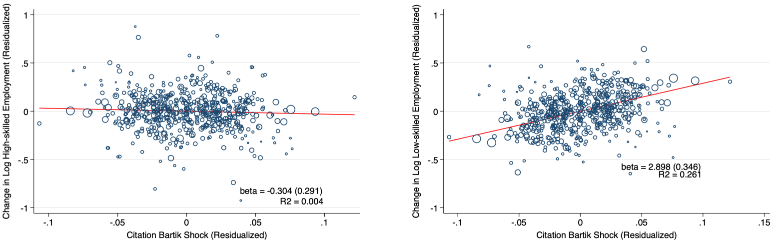



Figure 4: Effect of Citation Shock on Number of High- and Low-skilled Employment

Patents

# Descriptive Analysis

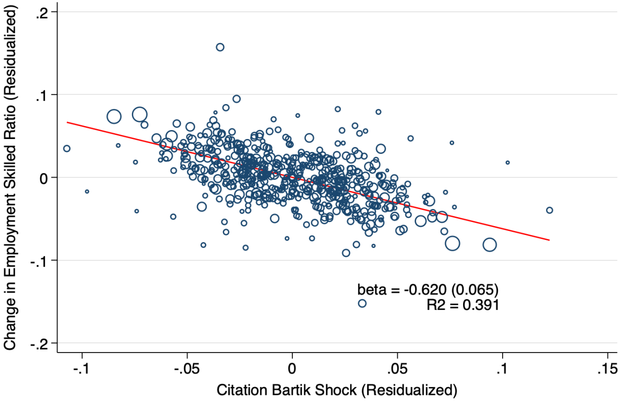



Figure 5: Citation Shock and Change in Skilled Ratio

Patents

# Descriptive Analysis

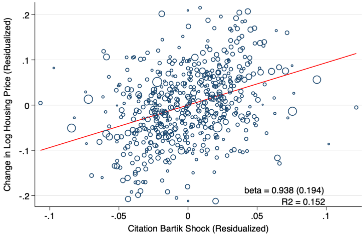



Figure 6: Effect of Citation Shock on Housing Price and Amenity

▸ Patents

# Descriptive Analysis: Regression Analysis

| VARIABLES | (1)<br>Δ Log Employment | (2)<br>Δ Log High-Skilled Employment | (3)<br>Δ Log Low-Skilled Employment | (4)<br>Δ Log High-skilled Migrants | (5)<br>Δ Log Low-skilled Migrants | (6)<br>Δ Employment Skilled Ratio |
|---|---|---|---|---|---|---|
| *Panel A: OLS* | | | | | | |
| Δ Log(Citation) | 0.0477 | -0.00261 | 0.0644 | 0.0545 | 0.0921 | -0.0134* |
| | (0.0325) | (0.0313) | (0.0394) | (0.0757) | (0.0570) | (0.00785) |
| *Panel B: Reduced Form* | | | | | | |
| Citation shock | 1.837*** | -0.304 | 2.898*** | 1.607*** | 3.051*** | -0.620*** |
| | (0.312) | (0.291) | (0.346) | (0.605) | (0.472) | (0.0654) |
| *Panel C: IV* | | | | | | |
| Δ Log(Citation) | 1.739** | -0.369 | 2.836** | 1.056* | 3.044** | -0.626** |
| | (0.772) | (0.357) | (1.240) | (0.565) | (1.441) | (0.276) |
| Year FE | X | X | X | X | X | X |
| City FE | X | X | X | X | X | X |

▸ Patents ▸ Validity of Bartik

# Descriptive Analysis: Regression Analysis

| VARIABLES | (1)<br>Δ Log<br>High-skilled Wage | (2)<br>Δ Log Low-skilled<br>Wage | (3)<br>Δ Log(Housing<br>Price) | (4)<br>Δ Amenity Index |
|---|---|---|---|---|
| *Panel A: OLS* | | | | |
| Δ Log(Citation) | -0.00957 | -0.0140 | -0.0239 | 0.0889 |
| | (0.0158) | (0.0167) | (0.0160) | (0.0630) |
| *Panel B: Reduced Form* | | | | |
| Citation shock | 0.727*** | 0.549*** | 0.938*** | 0.709 |
| | (0.121) | (0.105) | (0.194) | (0.462) |
| *Panel C: IV* | | | | |
| Δ Log(Citation) | 0.727** | 0.560* | 1.017 | 0.955 |
| | (0.352) | (0.292) | (0.623) | (0.688) |
| Year FE | X | X | X | X |
| City FE | X | X | X | X |

Patents

# Descriptive Statistics

- We have the following findings in our data
- As technology grew in China during the last decade:
  - Wages for both skills increased, low-skilled workers migrated more
  - Housing price increased
  - Amenity increased, but not in a very significant magnitude
- In general, we do not find positive sorting of migration resulting from the technology growth in China during 2005-2015
- This is totally different from previous findings in developed countries
- Why is this the case?
- Let's go to a structural model to explain these findings

# Model Setup

- $K$ cities in China, indexed by $k \in \{1, ..., K\}$
- Two sectors, $j \in \{a, na\}$
- Workers differ in home location $k_0$, hukou type $j_0$, and skill $e \in \{L, H\}$.
- Each worker $i$ first chooses the sector $j$, then chooses which city $k$ to live
- Non-ag hukou workers can only choose the non-ag sector
- Three markets: labor market, housing market, and amenity market
- Labor supply, housing supply, and amenity supply are endogenous in the model

# Labor Demand — Non-agricultural Sector

- Firms in the non-agricultural sector produce a homogeneous tradable good
- Using technology $A_{kt}$, high-skilled labor $H_{na,kt}$, low-skilled labor $L_{na,kt}$, capital $K_{na,kt}$, and machine $C_{na,kt}$

$$
\begin{aligned}
Y_{na,kt} &= z_{na,kt} N_{na,kt}^{\alpha} (\theta_{kt}^{K} K_{kt})^{1-\alpha} \\
N_{na,kt} &= (\theta_{kt}^{L} (L_{na,kt} + \omega C_{kt})^{\rho} + \theta_{kt}^{H} H_{na,kt}^{\rho})^{\frac{1}{\rho}} \\
C_{kt} &= f_C(A_{kt}) \\
\theta_{kt}^{K} &= f_K(A_{kt}) \\
\theta_{kt}^{L} &= f_L(A_{kt}, H_{na,kt}, L_{na,kt}) \\
\theta_{kt}^{H} &= f_L(A_{kt}, H_{na,kt}, L_{na,kt})
\end{aligned}
$$

- $N_{na,kt}^{\alpha_{na}}$ is a CES aggregator of labor
- $L$ and $C$ are substitutes
- $\theta$ is factor-augmenting productivity

# Labor Demand — Non-agricultural Sector

- This is a non-linear C-D style production function with many premitive parameters
- We can transform this to be a linear labor demand equation
- Using first order conditions and log linearization

# Labor Demand — Non-agricultural Sector

- By F.O.C. we have:

$$
\begin{aligned}
W^H_{na,kt} &= z_{na,kt}\alpha N^{\alpha-\rho}_{na,kt}(\theta^K_{kt}K_{kt})^{1-\alpha}H^{\rho-1}_{na,kt}\theta^H_{kt} \\
W^L_{na,kt} &= z_{na,kt}\alpha N^{\alpha-\rho}_{na,kt}(\theta^K_{kt}K_{kt})^{1-\alpha}(L_{na,kt}+\omega C_{kt})^{\rho-1}\theta^L_{kt} \\
\kappa_t &= z_{kt}N^{\alpha}_{na,kt}(\theta^K_{kt}K_{kt})^{-\alpha}(1-\alpha)\theta^K_{kt}
\end{aligned}
$$

# Labor Demand — Non-agricultural Sector

- We then log linearize this system and have labor demand functions as:

$$w^H_{na,kt} = \ln W^H_{na,kt} = d_{na,kt} + (1-\rho)\ln N_{na,kt} + (\rho - 1)\ln H_{na,kt} + \ln \theta^H_{kt}$$

$$w^L_{na,kt} = \ln W^L_{na,kt} = d_{na,kt} + (1-\rho)\ln N_{na,kt} + (\rho - 1)\ln(L_{na,kt} + \omega C_{kt}) + \ln \theta^L_{kt}$$

$$N_{na,kt} = (\theta^L_{kt}(L_{na,kt} + \omega C_{kt})^\rho + \theta^H_{kt} H^\rho_{na,kt})^{\frac{1}{\rho}}$$

$$d_{na,kt} = \ln\left( z^{1/\alpha}_{na,kt}\alpha \left(\frac{(1-\alpha)\theta^K_{kt}}{\kappa_t}\right)^{\frac{1-\alpha}{\alpha}}\right)$$

- Does this system look familiar?

# Labor Demand — Non-agricultural Sector

- We can write the wage as a additive separable function of workforces $H$, $L$ and technology $A$:

$$w^H_{na,kt} = g_{na,H}(A_{kt}, H_{na,kt}, L_{na,kt}) + d^H_{na,kt}$$
$$w^L_{na,kt} = g_{na,L}(A_{kt}, H_{na,kt}, L_{na,kt}) + d^L_{na,kt}$$

- $d$ is the structural residual, determined by various of deep parameters in the production function
- Then we can use a simple log linear function to approximate them

# Labor Demand — Non-agricultural Sector

- We further rewrite it a little bit:

$$w^H_{na,kt} = \gamma_{HA} A_{kt} + \gamma_{na,HH} \ln(H_{na,kt}) + \gamma_{na,HL} \ln(L_{na,kt}) + d^H_{na,kt}$$
$$w^L_{na,kt} = \gamma_{LA} A_{kt} + \gamma_{na,LH} \ln(H_{na,kt}) + \gamma_{na,LL} \ln(L_{na,kt}) + d^L_{na,kt}$$

- These are just two linear regressions!
- Log wage is $y$, log workforce numbers are $x$, error term $d$
- $A_{kt}$ is proxied by the patent shock

# Labor Demand — Non-agricultural Sector

- If we do not care so much about the details of the production function
- If you just want to model the wage/labor demand at equilibrium
- We can just estimate regression coefficents $\gamma$ using IV
- Instead of estimating the primitive parameters in the production function $\theta, z, \omega$
- This makes your life much much much easier
- But this is not feasible when you want to look deep into the production process
- Or if you want to run some counterfactuals by changing $z$ or other deep parameters

# Labor Demand - Agricultural Sector

- The production in the agricultural sector only involves high-skilled labor $H_{a,kt}$ and low-skilled labor $L_{a,kt}$
- $H$ and $L$ are perfect substitutes

$$Y_{a,kt} = z_{ag,kt}(N_{a,kt}^{\alpha_a})^{\eta}$$
$$N_{a,kt} = L_{a,kt} + H_{a,kt}$$

- Similarly, we can transform labor demand function in agricultural sector as:

$$w_{ag,kt}^{H} = w_{ag,kt}^{L} = \gamma_{ag} \ln(H_{ag,kt} + L_{ag,kt}) + d_{ag,kt}$$

# Labor Supply

- Four types of workers: ag local, ag migrant, na local, na migrant
- Two restrictions in location choice set
  - Workers with non-ag hukou do not choose ag sector
  - Workers with ag hukou do not choose ag sector in other locations
- Thus, workers with non-ag hukou choose among non-ag sectors in all locations in one step
- Workers with ag hukou make sequential decisions
  - Step 1: Choose between hometown ag sector and non-ag sector
  - Step 2: If choosing non-ag sector, which location to go

# Labor Supply: Location Choices

The utility of working in the non-agricultural sector in city $k$ and year $t$ is

$$V_{ikt} = \beta_e^w (w_{na,kt}^e - \zeta r_{kt}) + \beta_e^h WithinHometown_{ikt} + \beta_e^p WithinProvince_{ikt}$$
$$+ \beta^h hukou_{kt} + \beta_e^a a_{kt} + \nu_{kt}^e + \epsilon_{ikt}$$

- $w_{kt}^e$ is wage, $r_{kt}$ is housing price; $\zeta$ is expenditure share on housing
- $a_{kt}$ and $\nu_{kt}$ are the observed endogenous/unobserved exogenous amenity
- $hukou_{kt}$ is the hukou policy
- $WithinHometown_{ikt}$ and $WithinProvince_{ikt}$ capture migration cost/home bias
- $\epsilon_{ikt}$ is an i.i.d. shock with T1EV distribution
- Heterogeneous preference varies by worker's skill $e$

# Labor Supply: Location Choices

- The value of working in the non-agricultural sector for individual $i$ in year $t$ with hometown $k_0$ is

$$W_{ik_0t}^{na} = \max\{V_{i1t}^{k_0}, V_{i2t}^{k_0}, ...V_{iKt}^{k_0}\}$$

  which is the maximum value of working in the non-agricultural over all possible cities.

- Based on the property of T1EV distribution, we have:

$$E[W_{ik_0t}^{na}] = \max\{V_{i1t}^{k_0}, V_{i2t}^{k_0}, ...V_{iKt}^{k_0}\} = ln[\sum_{k \in K} exp(V_{ikt}^{k_0})]$$

# Labor Supply: Sector Choices

- For ag hukou workers, they have an additional first step sector choice
- The value of working in the agricultural sector in prefecture $k$ is

$$W_{it}^{a} = \alpha_{0k_0} + \alpha_1 w_{a,k_0 t}^{e} + \xi_{it}^{a}$$
$$W_{it}^{na} = E[W_{ik_0 t}^{na}] + \xi_{it}^{na}$$

- $\alpha_{0k}$ is a city-specific constant term
- $w_{a,kt}$ is the agricultural earnings
- $\xi_{it}$ is an i.i.d. shock with T1EV distribution
- The sector choice decision is

$$\max\{W_{it}^{na}, W_{it}^{a}\}$$

# Labor Supply: Sector Choices

- This is a very typical discrete choice model
- It is non-linear, but still we have many off-the-shelf tools to estimate it
- We will show later how to estimate this non-linear system using IV: BLP method

# Housing Supply

- Developers are price-takers and sell homogeneous houses at the marginal cost of production

$$R_{kt} = \iota_t \times MC(CC_{kt}, LC_{kt})$$

where $\iota_t$ is interest rate, $CC_{kt}$ is construction costs, and $LC_{kt}$ is land costs.

- The cost of land $LC_{kt}$ is a function of the aggregate demand for local goods

$$HD_{kt} = L_{na,kt} W^{L}_{na,kt} + H_{na,kt} W^{H}_{na,kt}$$

- The housing supply equation is

$$\begin{aligned} r_{kt} &= \ln(R_{kt}) = \ln(\iota_t) + \ln(CC_{kt}) + \gamma_k \ln(HD_{kt}) \\ \gamma_k &= \gamma^{hd}_1 + \gamma^{hd}_2 geo_k \end{aligned}$$

where $x^{geo}_k$ is the altitude that affects the elasticity of housing price with respect to local good demand.

# Amenity Supply

- Endogenous amenity depends on technology and skill ratio:

$$a_{kt} = \gamma_1^a A_{kt} + \gamma_2^a \ln\left(\frac{H_{na,kt}}{L_{na,kt}}\right) + \epsilon_{kt}^a$$

- Local amenities respond to the education of neighboring households (Bayer, Ferreira, and McMillan, 2007) and college employment ratios (Diamond, 2016).
- Patents/tech growth have a direct impact on amenities.

# Equilibrium Definition

Equilibrium in this model is defined by a set of working populations, wages, housing prices, and amenities such that

- The high-skill labor demand equals high-skill labor supply for both sectors and all cities.
- The low-skill labor demand equals low-skill labor supply for both sectors and all cities.
- Housing demand equals housing supply in the non-agricultural sector for all cities.
- Endogenous amenities demand equals endogenous amenity supply for both sectors and all cities.

# Estimation Results

- We log linearize and take the first difference of these equations
- Then use different Bartik IVs to estimate the parameters of the model
- In estimation, we find that:
  - Patent growth increases wages ▸ Estimation of Labor Demand
  - Patent growth increases housing prices ▸ Estimation of Housing Supply
  - Patent growth increases amenity ▸ Estimation of Amenity Supply
  - Low-skilled care more about wage; High-skilled care more about amenity ▸ Estimation of Labor Supply

The model fit is good ▸ Model Fit

# Estimation of Labor Demand in the Non-agricultural Sector

- We use first difference regression to estimate labor demand equations

$$\Delta w^{H}_{na,kt} = \gamma_{HA}\Delta A_{kt} + \gamma_{na,HH}\Delta \ln H_{na,kt} + \gamma_{na,HL}\Delta \ln L_{na,kt} + \Delta\epsilon^{H}_{na,kt}$$
$$\Delta w^{L}_{na,kt} = \gamma_{LA}\Delta A_{kt} + \gamma_{na,LH}\Delta \ln H_{na,kt} + \gamma_{na,LL}\Delta \ln L_{na,kt} + \Delta\epsilon^{L}_{na,kt}$$

- Need variation in labor supply uncorrelated with unobserved changes in local productivity
- Instrument for low- and high-skilled workforce: migrant Bartik
- Instrument for local patent shock: patent Bartik

Back

# Migrant Bartik

- National changes in the number of migrants at industry level $\times$ share of migrants in each industry in a location (Card, 2009):

$$\Delta B^e_{kt} = \sum_{\text{ind}} \left( Mig^e_{\text{ind}\ ,na,-k,t} - Mig^e_{\text{ind}\ ,na,-k,2005} \right) \frac{Mig^e_{ind,na,k,2005}}{Mig^e_{na,k,2005}}$$

- $Mig^e_{ind,na,-k,t}$: total number of skill $e$ migrants in industry ind in year $t$, excluding migrants in city $k$
- $Mig^e_{ind,na,k,2005}$: number of skill $e$ migrants in industry ind in city $k$ in year 2005.

Back

# Patent Bartik Shock

$$\Delta P_{kt} = \sum_{ind}(Patent_{ind,-k,t} - Patent_{ind,-k,2005})\frac{E_{ind,k,2005}}{E_{k,2005}}$$

- $Patent_{ind,-k,t}$ represents the log number of patent in industry *ind* in year $t$ in the country, excluding city $k$.
- $E_{ind,k,2005}$ and $E_{k,2005}$ measure the number of workers in industry *ind* in city $k$ in year 2005.
- Measure exogenous technology shocks
- Industry: two-digit, total 95 industries (54 with patent)

Back

# Estimation of Labor Demand in the Non-agricultural Sector

Table 1: Estimation of the Labor Demand in the Non-agricultural Sector

| VARIABLES | (1) Δ Log High-skilled Wage | (2) Δ Log Low-skilled Wage | (3) Δ Log High-skilled City Population | (4) Δ Log Low-skilled City Population | (5) Δ Log Citation | (6) Δ Log High-skilled Wage | (7) Δ Log Low-skilled Wage |
|---|---|---|---|---|---|---|---|
| Δ Log Citation | 0.0330*** | 0.0330** | | | | 1.128*** | 1.055*** |
| | (0.0118) | (0.0131) | | | | (0.330) | (0.313) |
| Δ Log High-skilled City Population | 0.0390 | 0.0408 | | | | -0.176 | 0.00360 |
| | (0.0241) | (0.0268) | | | | (0.520) | (0.494) |
| Δ Log Low-skilled City Population | -0.0601* | -0.0639* | | | | -0.691 | -0.838 |
| | (0.0310) | (0.0344) | | | | (0.782) | (0.743) |
| Citation shock | | | -0.208 | 0.0923 | 0.514** | | |
| | | | (0.134) | (0.106) | (0.236) | | |
| Migrant bartik for high-skill workers | | | 0.607*** | 0.387*** | 0.367 | | |
| | | | (0.141) | (0.112) | (0.248) | | |
| Migrant bartik for low-skill workers | | | -0.288 | 0.548*** | -0.259 | | |
| | | | (0.249) | (0.197) | (0.438) | | |
| Observations | 481 | 481 | 481 | 481 | 481 | 481 | 481 |
| R-squared | 0.025 | 0.021 | | | | | |
| Model | OLS | OLS | First stage | First stage | First stage | IV GMM | IV GMM |
| Cragg-Donald Wald F | | | | | | 3.355 | 3.355 |
| Sanderson-Windmeijer F | | | 14.51 | 14.42 | 10.73 | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Back Patents

# Estimation of Labor Demand in the Agricultural Sector

$$\Delta w_{a,kt} = \gamma_a \Delta \ln(H_{a,kt} + L_{a,kt}) + \Delta \epsilon_{a,kt}$$

- Count all workers in the rural area as agricultural workers.
- Instrument for agricultural employment: population with agricultural hukou

Back

# Estimation of Labor Demand in the Agricultural Sector

Table 2: Estimation of the Labor Demand in the Agricultural Sector

| VARIABLES | (1)<br>Δ Log Agr Income | (2)<br>Δ Log Agr Employ-ment | (3)<br>Δ Log Agr Income |
|---|---|---|---|
| Δ log agricultural employment | -0.0423* | | -0.166*** |
| | (0.0238) | | (0.0341) |
| Δ log agricultural population | | 1.166*** | |
| | | (0.0518) | |
| Observations | 481 | 481 | 481 |
| R-squared | 0.007 | | |
| Model | OLS | First stage | IV GMM |
| Cragg-Donald Wald F | | | 506.4 |
| Sanderson-Windmeijer F | | 506.4 | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Back

# Estimation of the Housing Market

$$\Delta r_{kt} = \left[\gamma_1^{hd} + \gamma_2^{hd} \times \ln(Slope_k)\right] \Delta \ln(HD_{kt}) + \Delta\epsilon_{kt}^r$$

- Need variation in housing demand unrelated to changes in unobserved factors driving housing prices $\Delta \ln(CC_{kt})$
- Instrument for housing demand: wage Bartik

Back

# Wage Bartik

- We interact cross-sectional differences in industrial employment composition with national changes in industry wage levels to construct the wage Bartik following (Diamond, 2016):

$$\Delta W_{kt}^{H} = \sum_{ind} (w_{ind,na,-k,t}^{H} - w_{ind,na,-k,2005}^{H}) \frac{H_{ind,na,k,2005}}{H_{na,k,2005}}$$

$$\Delta W_{kt}^{L} = \sum_{ind} (w_{ind,na,-k,t}^{L} - w_{ind,na,-k,2005}^{L}) \frac{L_{ind,na,k,2005}}{L_{na,k,2005}}$$

- $w_{ind,na,-k,t}^{e}$: log wage of high/low skill workers in industry *ind* in year $t$, excluding city $k$
- $H_{ind,na,k,2005}$ and $L_{ind,na,k,2005}$: high/low skill workers in industry *ind* in city $k$ in the urban area in 2005

Back

# Estimation of the Housing Market

Table 3: Estimation of the Housing Market

| VARIABLES | (1) Δ Log(Housing Price) | (2) Δ Log Housing Demand | (3) Δ Log Housing Demand * Geo | (4) Δ Log(Housing Price) |
|---|---|---|---|---|
| Δ Log housing demand | 0.147*** | | | 0.376*** |
| | (0.0201) | | | (0.0443) |
| Δ Log housing demand * Log slope | 0.0268*** | | | 0.0206** |
| | (0.00798) | | | (0.0100) |
| Wage Bartik IV for High-skilled Workers | | 2.131*** | -1.963*** | |
| | | (0.381) | (0.516) | |
| Wage Bartik IV for Low-skilled Workers | | -1.169*** | 1.612*** | |
| | | (0.387) | (0.525) | |
| Wage Bartik IV for High-skilled Workers * Log slope | | -0.917*** | 1.720*** | |
| | | (0.237) | (0.321) | |
| Wage Bartik IV for Low-skilled Workers * Log slope | | 0.954*** | -0.489 | |
| | | (0.237) | (0.322) | |
| Observations | 481 | 481 | 481 | 481 |
| R-squared | 0.147 | | | |
| Model | OLS | First stage | First stage | IV GMM |
| Cragg-Donald Wald F | | | | 40.65 |
| Sanderson-Windmeijer F | | 622.1 | 54.93 | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

▸ Back

# Estimation of the Amenity Market

$$\Delta a_{kt} = \gamma_1^a \Delta A_{kt} + \gamma_2^a \Delta \ln \left( \frac{H_{na,kt}}{L_{na,kt}} \right) + \Delta \epsilon_{kt}^a$$

- Instrument for changes in the skilled ratio: employment Bartik for low- and high-skilled workers
- Instrument for patent shocks: patent Bartik

Back

# Employment Bartik

$$\Delta E^H_{kt} = \sum_{ind} (H_{ind,na,-k,t} - H_{ind,na,-k,2005}) \frac{H_{ind,na,k,2005}}{H_{na,k,2005}}$$

$$\Delta E^L_{kt} = \sum_{ind} (L_{ind,na,-k,t} - L_{ind,na,-k,2005}) \frac{L_{ind,na,k,2005}}{L_{na,k,2005}}$$

- $H_{ind,na,-k,t}$ represents the log number of high-skilled worker in industry *ind* in year $t$ in the country, excluding city $k$.
- $L_{ind,na,-k,t}$ represents the log number of low-skilled worker in industry *ind* in year $t$ in the country, excluding city $k$.
- $H_{na,-k,t}$, $L_{na,-k,t}$ follow the same definitions as in the construction of wage bartik.

▸ Back

# Estimation of the Amenity Market

Table 4: Estimation of the Amenity Market

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | Δ Amenity Index | Δ Log High-skilled City Population Ratio | Δ Log Citation | Δ Amenity Index |
| Δ Log Citation (Flow, IPA) | 0.152*** | | | 1.109*** |
| | (0.0350) | | | (0.315) |
| Δ High-skilled City Population Ratio | 0.397 | | | 4.714** |
| | (0.396) | | | (2.397) |
| Citation shock | | -0.0807*** | 0.374 | |
| | | (0.0211) | (0.246) | |
| Wage Bartik IV for high-skilled workers | | 0.0757 | -1.055 | |
| | | (0.0576) | (0.671) | |
| Wage Bartik IV for low-skilled workers | | 0.0792 | 1.579** | |
| | | (0.0611) | (0.713) | |
| Observations | 481 | 481 | 481 | 481 |
| R-squared | 0.040 | | | |
| Model | OLS | First stage | First stage | IV GMM |
| Cragg-Donald Wald F | | | | 5.038 |
| Sanderson-Windmeijer F | | 13.15 | 7.846 | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

▸ Back ▸ Patents ▸ Sub-indices

# Estimation of the Amenity Market

Table 5: Estimation of Amenity Market (Sub-indices)

| VARIABLES | (1)<br>Δ Infrastructure Index | (2)<br>Δ Environment Index | (3)<br>Δ Health Index | (4)<br>Δ Education Index |
|---|---|---|---|---|
| Δ Log(Citation) | 0.968*** | -0.782*** | 0.797** | -0.0619 |
| | (0.289) | (0.235) | (0.317) | (0.0908) |
| Δ High-skilled Ratio | 4.777** | -4.425** | 1.292 | 2.696*** |
| | (2.207) | (1.797) | (2.418) | (0.693) |
| Constant | -0.903** | 1.420*** | -0.879* | 0.153 |
| | (0.418) | (0.340) | (0.458) | (0.131) |
| Observations | 481 | 481 | 481 | 481 |
| Model | IV GMM | IV GMM | IV GMM | IV GMM |
| Kleibergen-Paap rk LM | 14.6 | 14.6 | 14.6 | 14.6 |
| Cragg-Donald Wald F | 4.978 | 4.978 | 4.978 | 4.978 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Back

# Estimation of the Amenity Market

Table 6: Estimation of Amenity Market (Robustness)

| VARIABLES | (1)<br>Δ Amenity Index | (2)<br>Δ Log High-skilled Ratio | (3)<br>Δ Log Citation | (4)<br>Δ Amenity Index |
|---|---|---|---|---|
| Δ Log Citation | 0.137*** | | | 1.036*** |
| | (0.0349) | | | (0.368) |
| Δ High-skilled Employment Ratio | 0.502 | | | 4.950** |
| | (0.394) | | | (2.400) |
| Δ Log number of college students | 0.266*** | -0.0232** | 0.168 | 0.0930 |
| | (0.0799) | (0.00903) | (0.106) | (0.182) |
| Citation Shock | | -0.0720*** | 0.306 | |
| | | (0.0214) | (0.251) | |
| Wage Bartik IV for High-skilled Workers | | 0.0686 | -1.012 | |
| | | (0.0574) | (0.672) | |
| Wage Bartik IV for Low-skilled Workers | | 0.0894 | 1.511** | |
| | | (0.0611) | (0.715) | |
| Constant | 0.406*** | 0.0288 | 0.696*** | -1.147** |
| | (0.0781) | (0.0207) | (0.243) | (0.467) |
| Observations | 479 | 479 | 479 | 479 |
| R-squared | 0.062 | | | |
| Model | OLS | First stage | First stage | IV GMM |
| Sanderson-Windmeijer F | | 9.432 | 5.386 | |

Back

# Labor Supply: Location Choices

- We have the following utility function to estimate:

$$V_{ikt} = \beta_e^w (w_{na,kt}^e - \zeta r_{kt}) + \beta_e^h WithinHometown_{ikt} + \beta_e^p WithinProvince_{ikt} + \beta_e^h hukou_{kt} + \beta_e^a a_{kt} + \nu_{kt}^e + \epsilon_{ikt} \tag{1}$$

- With T1EV distribution, it is a typical Logit model
- We can always use MLE to estimate it
- However, we all know that wage, rent, amenity are all endogenous in the model
- Migration choices of people will affect them

Back

# Endogeneity in DCM

- We must find a way to solve this endogeneity issue
- In linear regression, we usually use IV
- Therefore, we have to find a way to apply IV in discrete choice model
- Traditional linear IV method did not work in this case
- We introduce two methods today: BLP and Control Function

Back

# Endogeneity in DCM: 1. BLP

- The first method is called BLP, introduced in Berry, Levinsohn, and Pakes (1995)
- It first transform the endogeneity issue in a nonlinear model to a linear one
- Then we can use well-developed linear IV method to solve it

# Endogeneity in DCM: 1. BLP

- Assume we have the car buying problem
- There are M markets with $J_m$ options (brands) in each market
- Utility for consumer $n$ in market $j$ to choose brand $m$ is:

$$U_{njm} = V(p_{jm}, x_{jm}, s_n, \beta_n) + \xi_{jm} + \epsilon_{njm}$$

- $p_{jm}$ price; $s_n$ personal attributes; $x_{jm}$ product attributes; $\xi_{jm}$ unobserved product attributes; $\epsilon_{njm}$ i.i.d. T1EV shock
- Price is correlated with unobserved brand attributes $\xi_{jm} \not\perp p_{mj}$

# Endogeneity in DCM: 1. BLP

- Important feature of BLP: endogeneity comes from market-product level $\xi_{jm}$
- That is why we can use it in Diamond model
- Migration choice is individual level
- But endogeneity comes from destination location level

# Endogeneity in DCM: 1. BLP

- The idea of BLP employs a two-step approach
- First, add in a product-market level FE, absorb $\xi_{jm}$
- Estimate the equation with fixed effect
- Second, open the box of product-market level FE, estimate the remaining parameters

# Endogeneity in DCM: 1. BLP

- We can decompose the observed utility value into

$$V_{njm} = \underbrace{\bar{V}(p_{jm}, x_{jm}, \bar{\beta})}_{\text{varies only over product-market}} + \underbrace{\tilde{V}(p_{jm}, x_{jm}, s_n, \tilde{\beta}_n)}_{\text{varies also over consumer}}$$

- Then we have the utility to be

$$U_{njm} = \underbrace{[\bar{V}(p_{jm}, x_{jm}, \bar{\beta}) + \xi_{jm}]}_{\text{a product-market level fixed effect}} + \tilde{V}(p_{jm}, x_{jm}, s_n, \tilde{\beta}_n) + \epsilon_{njm}$$

- We just combine all terms varying only at product-market level together

# Endogeneity in DCM: 1. BLP

- We define product-market level FE as:

$$\delta_{jm} = \bar{V}(p_{jm}, x_{jm}, \bar{\beta}) + \xi_{jm} \tag{2}$$

$$U_{njm} = \delta_{jm} + \tilde{V}(p_{jm}, x_{jm}, s_n, \tilde{\beta}_n) + \epsilon_{njm} \tag{3}$$

- Equation (3) does not entail any endogeneity
- Step 1: We run a Logit model with *jm* level FE to estimate parameters $\tilde{\beta}$
- Step 2: We get estimates of $\delta_{jm}$ in step 1, and run IV regression for equation and get $\bar{\beta}$ (2)

# Endogeneity in DCM: 1. BLP

The essence of BLP

- We cannot run IV regression directly in DCM
- We first pack all terms at the level where endogeneity happens into FE
- Then we estimate a DCM with these FEs
- We have estimated FEs, then unpack it and run linear IV regression
- Transform non-linear IV to be linear IV
- BLP tells you how to use an IV in a DCM, but only in a specific model structure with nested endogeneous variables.

# Endogeneity in DCM: 2. Control Function

- BLP is not always feasible (error structure...)
- The algorithm of estimating BLP is sometimes complicated
- If the dimension of the fixed effect is too high, you have to use some contraction method
- Highly recommend you to read BLP part in Train's book (or better, BLP 1993)
- The second important non-linear IV approach is Control Function (CF)

# Endogeneity in DCM: 2. Control Function

- The utility of consumer $n$ buying product $j$ is:

$$U_{nj} = V(y_{nj}, x_{nj}, \beta_n) + \epsilon_{nj}$$

- $y_{nj}$ is endogenous, $y_{nj} \not\perp \epsilon_{nj}$
- We assume that there is an instrument $z_{nj}$, related with $y_{nj}$ by first stage:

$$y_{nj} = W(z_{nj}, \gamma) + \mu_{nj} \tag{4}$$

- Assume that $\epsilon_{nj}, \mu_{nj} \perp z_{nj}$, $\epsilon_{nj} \not\perp \mu_{nj}$
- $\epsilon_{nj} \not\perp \mu_{nj}$ implies that $y_{nj}$ and $\epsilon_{nj}$ are correlated

# Endogeneity in DCM: 2. Control Function

- Therefore, $\mu$ is the source of the endogeneity
- We want to to some extent control it
- We can do a CEF decomposition (given $\mu_{nj}$) for $\epsilon_{nj}$:

$$\epsilon_{nj} = \underbrace{E(\epsilon_{nj}|\mu_{nj})}_{CF(\mu_{nj},\lambda)} + \tilde{\epsilon}_{nj}$$

- By construction: $\tilde{\epsilon}_{nj} \perp \mu_{nj}$
- Thus, we have $\tilde{\epsilon}_{nj} \perp y_{nj}$ ($y$ is correlated with $\epsilon$ only through $\mu$)
- We call $CF(\mu_{nj}, \lambda)$ a control function, where $\lambda$ is some parameter

# Endogeneity in DCM: 2. Control Function

- Then we can rewrite the utility function as

$$U_{nj} = V(y_{nj}, x_{nj}, \beta_n) + CF(\mu_{nj}, \lambda) + \tilde{\epsilon}_{nj} \tag{5}$$

- Step 1: Estimate first stage equation (4), get residual of the first stage $\hat{\mu}$
- Step 2: Plug $\hat{\mu}$ in the CF (5)
- Step 3: Estimate equation (5) using simple Logit
- In step 2, we need to assume a functional form for CF
- Usually we can choose flexible non-parametric form (e.g. high-order polynomials)

# Endogeneity in DCM: 2. Control Function

- The logic of CF approach is as follows:
  - We know that instrument $z$ is not correlated with the error $\epsilon$
  - Thus, endogenous variable $y$ correlates with $\epsilon$ only through first stage error $\mu$, but not IV $z$
  - Then by controlling the correlated parts of $\mu$ and $\epsilon$, we can eliminate the correlation of $y$ and $\epsilon$
- CF is a pretty general method
- But it requires you to set a function form for CF

# Labor Supply: Location Choices

- Now we go back to our paper
- We choose BLP method to tackle the endogeneity issue
- Our utility function has a nested error structure
- Choice is made at individual level
- Endogeneity happens at location level: wages, housing prices, amenities, and Hukou policies
- So we can estimate the model by adding location fixed effect, then regress this fixed effect on endogeneous variables with IV

# Labor Supply: Location Choices

BLP model:

$$V_{ikt}^{k_0} = \beta_h^e WithinHometown_{ikt} + \beta_p^e WithinProvince_{ikt} + \delta_{kt}^e + \epsilon_{ikt}$$
$$\delta_{kt}^e = \beta_w^e (w_{kt}^e - \zeta r_{kt}) + \beta_h^e \text{Hukou} + \beta_a^e a_{kt} + \nu_{kt}^e$$

- First step: maximum likelihood, get $\delta_{kt}^e$.
- Second step:

$$\Delta\delta_{kt}^e = \beta_w^e (\Delta w_{kt}^e - \zeta \Delta r_{kt}) + \beta_h^e \Delta\text{Hukou} + \beta_a^e \Delta a_{kt} + \Delta\epsilon_{kt}^e$$

  - $\zeta$: expenditure share on housing: mortgage share 40%, rent share 30%, set $\zeta = 0.35$; alternatively, set $\zeta = 0.62$ following Diamond (2016).
  - Instruments: choose from a large set of potential IVs using Lasso (Chernozhukov et al. 2018).

Back

Table 7: Labor Supply Estimation (Location Choice, BLP first stage)

| Worker Type | Year | Within Hometown (East) | Within Hometown (Middle) | Within Hometown (West) | Within Hometown (Northeast) | Within Province |
|---|---|---|---|---|---|---|
| Low-Skilled Worker | 2005 | 5.641*** | 6.084*** | 5.906*** | 6.317*** | 3.054*** |
| | | (0.021) | (0.027) | (0.039) | (0.013) | (0.008) |
| Low-Skilled Worker | 2010 | 4.743*** | 5.466*** | 4.639*** | 6.099*** | 2.993*** |
| | | (0.02) | (0.017) | (0.033) | (0.034) | (0.031) |
| Low-Skilled Worker | 2015 | 4.732*** | 5.118*** | 4.525*** | 5.767*** | 4.035*** |
| | | (0.025) | (0.034) | (0.031) | (0.014) | (0.01) |
| High-Skilled Worker | 2005 | 5.922*** | 6.635*** | 6.251*** | 6.935*** | 2.689*** |
| | | (0.03) | (0.033) | (0.04) | (0.025) | (0.01) |
| High-Skilled Worker | 2010 | 5.113*** | 5.548*** | 5.122*** | 5.874*** | 3.212*** |
| | | (0.026) | (0.039) | (0.04) | (0.016) | (0.011) |
| High-Skilled Worker | 2015 | 4.795*** | 5.115*** | 4.757*** | 5.672*** | 3.818*** |
| | | (0.027) | (0.001) | (0.03) | (0.019) | (0.012) |

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Back

# Second Step: Lasso for each endogenous variable

| Variables | Wage - 0.35 * Price (Low-Skilled) | Wage - 0.35 * Price (High-Skilled) | Amenity | Hukou Index |
|---|---|---|---|---|
| employment bartik | ✓ | ✓ | | ✓ |
| migrant bartik | | | ✓ | ✓ |
| wage bartik | | | ✓ | ✓ |
| export bartik | ✓ | ✓ | | ✓ |
| citation bartik | | ✓ | | |
| patent bartik | ✓ | ✓ | ✓ | ✓ |
| patent iv | ✓ | ✓ | | ✓ |
| import bartik | ✓ | | | ✓ |
| net export bartik | | ✓ | | |
| openness shock | ✓ | ✓ | ✓ | ✓ |
| robot bartik | | ✓ | | ✓ |
| geographics | ✓ | ✓ | ✓ | ✓ |
| land supply | ✓ | ✓ | ✓ | ✓ |
| export bartik × geographics | | | | ✓ |
| export bartik × land supply | ✓ | | | ✓ |
| openness shock × land supply | | | ✓ | ✓ |
| land supply × geographics | ✓ | ✓ | ✓ | ✓ |
| migrant bartik × robot bartik | | | | ✓ |
| wage bartik × geographics | | | ✓ | ✓ |
| patent bartik × geographics | ✓ | | | ✓ |

Back

# Second Step: Lasso for each endogenous variable

Table 8: First stage statistics of BLP second stage IV regression

| VARIABLES | (1)<br>First Stage R-squared | (2)<br>First Stage F-value |
|---|---|---|
| Δ Log(HS Wage-0.35*Price) | 0.6684 | 99.45 |
| Δ Log(LS Wage-0.35*Price) | 0.5096 | 46.10 |
| Δ Amenity Index | 0.1702 | 16.57 |
| Δ Hukou Index | 0.2569 | 9.86 |

Back

# Labor Supply: Location Choices

Table 9: Labor Supply Estimation (Location Choice, BLP second stage)

| VARIABLES | (1)<br>$\Delta\delta_{high}$ | (2)<br>$\Delta\delta_{low}$ | (3)<br>$\Delta\delta_{high}$ | (4)<br>$\Delta\delta_{low}$ |
|---|---|---|---|---|
| Δ Log(High-skilled Wage - 0.35*Price) | 0.0850 | | 0.0525 | |
| | (0.160) | | [0.462] | |
| Δ Log(Low-skilled Wage - 0.35*Price) | | 0.334 | | 2.620*** |
| | | (0.237) | | [0.784] |
| Δ Amenity Index | 0.0608 | 0.110* | 0.996*** | 0.148 |
| | (0.0504) | (0.0575) | [0.288] | [0.485] |
| Δ Hukou Index | -0.0635*** | -0.0345 | -0.399*** | -0.579*** |
| | (0.0203) | (0.0217) | [0.0687] | [0.0770] |
| Observations | 480 | 480 | 451 | 451 |
| R-squared | 0.020 | 0.029 | 0.080 | 0.156 |
| Model | OLS | OLS | IV | IV |

Back

# Labor Supply: Location Choices

Table 10: Labor Supply Estimation (Location Choice, BLP second stage)

| VARIABLES | (1)<br>$\Delta\delta_{high}$ | (2)<br>$\Delta\delta_{low}$ | (3)<br>$\Delta\delta_{high}$ | (4)<br>$\Delta\delta_{low}$ |
|---|---|---|---|---|
| $\Delta$ Log(High-skilled Wage - 0.35*Price) | 0.0873 | | 0.546 | |
| | (0.160) | | [0.413] | |
| $\Delta$ Log(Low-skilled Wage - 0.29*Price) | | 0.344 | | 2.367*** |
| | | (0.238) | | [0.698] |
| $\Delta$ Amenity Index | 0.0612 | 0.110* | 0.370*** | 0.153 |
| | (0.0504) | (0.0576) | [0.140] | [0.277] |
| $\Delta$ Hukou Index | -0.0645*** | -0.0358* | -0.345*** | -0.537*** |
| | (0.0203) | (0.0217) | [0.0705] | [0.0784] |
| Observations | 480 | 480 | 451 | 451 |
| R-squared | 0.020 | 0.030 | 0.072 | 0.148 |
| Model | OLS | OLS | IV | IV |

Back

# Labor Supply: Sector Choices

Table 11: Labor Supply Estimation (Sector Choice)

| Year | Skill | Coefficient on $w^a$ |
|---|---|---|
| 2005 | Low-Skilled | 1.377*** |
| | | (0.006) |
| 2010 | Low-Skilled | 1.269*** |
| | | (0.006) |
| 2015 | Low-Skilled | 1.454*** |
| | | (0.011) |
| 2005 | High-Skilled | 1.175*** |
| | | (0.045) |
| 2010 | High-Skilled | 1.206*** |
| | | (0.015) |
| 2015 | High-Skilled | 1.401*** |
| | | (0.019) |

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Back

# Mechanism

- Low-skilled workers:
  - Patent $\uparrow \Rightarrow$ wage $\uparrow \Rightarrow$ migrants $\uparrow\uparrow$ (Important)
  - Patent $\uparrow \Rightarrow$ amenity $\uparrow \Rightarrow$ migrants $\uparrow$ (Less Important)
- High-skilled workers:
  - Patent $\uparrow \Rightarrow$ wage $\uparrow \Rightarrow$ migrants $\uparrow$ (Less Important)
  - Patent $\uparrow \Rightarrow$ amenity $\uparrow \Rightarrow$ migrants $\uparrow$ (Important)
  - Patent $\uparrow \Rightarrow$ skilled-ratio $\downarrow \Rightarrow$ amenity $\downarrow \Rightarrow$ migrants $\downarrow$ (Important)

# Solving the Equilibrium: Algorithm

- The endogenous variables of city $k$ in year $t$ include $\mathbf{\Delta_0} = \{\mathbf{H_0}, \mathbf{L_0}, \mathbf{W_0}, \mathbf{R_0}, \mathbf{x_0}\}$
- Let $N_{k_0}^{a,s}$ and $N_{k_0}^{na,s}$ be the number of skill $s$ agricultural and non-agricultural hukou workers from hometown city $k_0$
- Let $q$ denote the iteration time
- Within each iteration, we use $\hat{var}$ to denote the temporary updating result of some variable $var$
- At the beginning of the q-th iteration, we have $\mathbf{\Delta_{q-1}}$.

# Solving the Equilibrium: Algorithm

- Step 1: update workers' utility values using endogenous variables derived from the last iteration $(q-1)$

$$\hat{\delta}_k^e = \beta_e^w(w_{na,k|q-1}^e - \zeta r_{k|q-1}) + \beta_e^h hukou_{k|q-1} + \beta_e^a a_{k|q-1} + \nu_k^e$$

$$\hat{V}_{ik}^{k_0} = \beta_e^h \text{ WithinHometown }_{ik|q-1}\beta_e^p \text{ WithinProvince }_{ik|q-1} + \hat{\delta}_k^e$$

$$E[\hat{W}_{ik_0}^{na}] = ln[\sum_{k\in K} exp(\hat{V}_{ik}^{k_0})]$$

# Solving the Equilibrium: Algorithm

- Step 2: update migration flows using the logit-form migration equations

$$\hat{H}_k^{na} = \sum_{i \in N^{na,H}} \frac{\exp(\hat{V}_{ik})}{\sum_r^K \exp(\hat{V}_{ir})} + \sum_{i \in N^{a,H}} \frac{exp(\hat{W}_{ik_0}^{na})}{exp(\hat{W}_i^a) + exp(\hat{W}_{ik_0}^{na})} \cdot \frac{\exp(\hat{V}_{ik})}{\sum_r^K \exp(\hat{V}_{ir})}$$

$$\hat{H}_k^{a} = \sum_{i \in N_k^{a,H}} \frac{exp(\hat{W}_i^a)}{exp(\hat{W}_i^a) + exp(\hat{W}_{ik}^{na})}$$

$$\hat{L}_{kt}^{na} = \sum_{i \in N^{na,L}} \frac{\exp(\hat{V}_{ik})}{\sum_r^K \exp(\hat{V}_{ir})} + \sum_{i \in N^{a,L}} \frac{exp(\hat{W}_{ik_0}^{na})}{exp(\hat{W}_i^a) + exp(\hat{W}_{ik_0}^{na})} \cdot \frac{\exp(\hat{V}_{ik})}{\sum_r^K \exp(\hat{V}_{ir})}$$

$$\hat{L}_{kt}^{a} = \sum_{i \in N_k^{a,L}} \frac{exp(\hat{W}_i^a)}{exp(\hat{W}_i^a) + exp(\hat{W}_{ik}^{na})}$$

# Solving the Equilibrium: Algorithm

- Step 3: update wages in each city using the wage equilibrium equation

$$\hat{w}^H_{ag,k} = \hat{w}^L_{ag,k} = \beta^1_0 + \gamma_{ag} \ln(\hat{H}_{ag,k} + \hat{L}_{ag,k}) + \epsilon_1$$
$$\hat{w}^H_{na,k} = \beta^2_0 + \gamma_{HA} A_k + \gamma_{na,LH} \ln \hat{H}_{na,k} + \gamma_{na,LL} \ln \hat{L}_{na,k} + \epsilon_2$$
$$\hat{w}^L_{na,k} = \beta^3_0 + \gamma_{LA} A_k + \gamma_{na,LH} \ln \hat{H}_{na,k} + \gamma_{na,LL} \ln \hat{L}_{na,k} + \epsilon_3$$

# Solving the Equilibrium: Algorithm

- Step 4: update the housing price in each city using housing equilibrium equation

$$\hat{r}_k = \beta_0^4 + \left[\gamma_1^{hd} + \gamma_2^{hd} \times \ln(Slope_k)\right] \ln(\hat{L}_k^{na} e^{\hat{w}_{na,k}^L} + \hat{H}_k^{na} e^{\hat{w}_{na,k}^H}) + \epsilon_4$$

- Step 5: update the amenity in each city using the amenity determination equation

$$\hat{a}_k = \beta_0^5 + \gamma_1^a A_k + \gamma_2^a \ln\left(\frac{\hat{H}_k^{na}}{\hat{L}_k^{na}}\right) + \epsilon_5$$

# Solving the Equilibrium: Algorithm

- Having these predicted values of the endogenous variables, we use the following updating rule to get the values of all variables for the next iteration:

$$\mathbf{\Delta_q} = \zeta \mathbf{\Delta_{q-1}} + (1-\zeta)\mathbf{\hat{\Delta}_{q-1}} \tag{6}$$

- $0 < \zeta < 1$, iterate until convergence is achieved

# Counterfactual 1: Patents Reduce to 2005 Level

- Change patent to the 2005 level
- Simulate labor supply, wages, housing price, and amenities in the general equilibrium framework

# Counterfactual 1: Patents Reduce to 2005 Level

Table 12: Eliminating Innovation Growth: Patent Citation Change in Log Points

| | Mean | Std Dev | Max | Min |
|---|---|---|---|---|
| National | -1.766 | 1.059 | 1.792 | -5.857 |
| Eastern Region | -1.879 | 0.706 | -0.231 | -3.734 |
| Middle Region | -2.128 | 1.017 | 0.274 | -5.857 |
| Northeastern Region | -0.646 | 0.879 | 1.792 | -1.861 |
| Western Region | -1.747 | 1.213 | 1.704 | -5.412 |

# Counterfactual 1: Patents Reduce to 2005 Level

Table 13: Eliminating Innovation Growth: Migration and Population Changes

| | Original Eq | Counterfactual | Change |
|---|---|---|---|
| **Panel A. Migration across Prefectures** | | | |
| Total migration | 27410758 | 21240364 | -22.51% |
| High-skilled migration | 5956344 | 6420181 | 7.79% |
| Low-skilled migration | 21454414 | 14820184 | -30.92% |
| **Panel B. Population** | | | |
| High-skilled pop in East | 25794098 | 26002678 | 0.81% |
| High-skilled pop in Middle | 9744213 | 9345408 | -4.09% |
| High-skilled pop in NE | 4425533 | 4634316 | 4.72% |
| High-skilled pop in West | 7039878 | 7021317 | -0.26% |
| Low-skilled pop in East | 95413904 | 92231360 | -3.34% |
| Low-skilled pop in Middle | 71139016 | 72099648 | 1.35% |
| Low-skilled pop in NE | 20001966 | 21833880 | 9.16% |
| Low-skilled pop in West | 42062284 | 42452280 | 0.93% |
| **Panel C. Urban Skill Ratio** | | | |
| Skill ratio in Urban East | 0.377 | 0.487 | 29.18% |
| Skill ratio in Urban Middle | 0.323 | 0.453 | 40.25% |
| Skill ratio in Urban NE | 0.359 | 0.342 | -4.74% |
| Skill ratio in Urban West | 0.349 | 0.447 | 28.08% |

Migrations decline dramatically for low-skilled workers.

# Counterfactual 1: Patents Reduce to 2005 Level

Table 14: Eliminating Innovation Growth: Wage Changes

| Skill | Sector | Region | Original Eq | Counterfactual | Change |
|---|---|---|---|---|---|
| Average Wage of Low-skilled | Agr | East | 15495 | 14863 | -4.08% |
| | | Middle | 11726 | 11384 | -2.92% |
| | | Northeast | 12046 | 11882 | -1.36% |
| | | West | 10460 | 10211 | -2.38% |
| | Non-agr | East | 51550 | 13335 | -74.13% |
| | | Middle | 43564 | 11175 | -74.35% |
| | | Northeast | 43199 | 22680 | -47.50% |
| | | West | 46864 | 14268 | -69.56% |
| Average Wage of High-skilled | Agr | East | 15495 | 14863 | -4.08% |
| | | Middle | 11726 | 11384 | -2.92% |
| | | Northeast | 12046 | 11882 | -1.36% |
| | | West | 10460 | 10211 | -2.38% |
| | Non-agr | East | 64749 | 14052 | -78.30% |
| | | Middle | 51148 | 10826 | -78.83% |
| | | Northeast | 50505 | 26913 | -46.71% |
| | | West | 57480 | 15618 | -72.83% |

Wages for both skills decline in similar magnitude.

# Counterfactual 1: Patents Reduce to 2005 Level

Table 15: Eliminating Innovation Growth: Housing Price and Amenity Changes

| | Original Eq | Counterfactual | Change |
|---|---|---|---|
| **Panel A. Housing Price** | | | |
| Average Housing Price in East | 7268.85 | 3544.92 | -51.23% |
| Average Housing Price in Middle | 4360.26 | 1967.82 | -54.87% |
| Average Housing Price in Northeast | 4327.30 | 3470.06 | -19.81% |
| Average Housing Price in West | 4691.29 | 2484.91 | -47.03% |
| **Panel B. Amenity** | | | |
| Average Amenity in East | 2.685 | 1.089 | -59.44% |
| Average Amenity in Middle | 2.564 | 0.856 | -66.63% |
| Average Amenity in Northeast | 2.823 | 2.140 | -24.18% |
| Average Amenity in West | 2.861 | 1.395 | -51.22% |

Amenity and housing price decline dramatically.

# Counterfactual 1: Patents Reduce to 2005 Level

Table 16: Eliminating Innovation Growth: Willingness to Pay

| Skill | Hukou Type | Region | WTP |
|---|---|---|---|
| Average WTP of Low-skilled | Agr | East | -0.095 |
| | | Middle | -0.080 |
| | | Northeast | -0.043 |
| | | West | -0.106 |
| | Non-agr | East | -1.501 |
| | | Middle | -1.485 |
| | | Northeast | -0.874 |
| | | West | -1.408 |
| Average WTP of High-skilled | Agr | East | -0.164 |
| | | Middle | -0.150 |
| | | Northeast | -0.110 |
| | | West | -0.150 |
| | Non-agr | East | -0.535 |
| | | Middle | -0.584 |
| | | Northeast | -0.349 |
| | | West | -0.532 |

Eliminating innovation harms all groups of workers, especially workers with non-agr hukou.

# Counterfactual 1: Patents Reduce to 2005 Level

Table 17: Eliminating Innovation Growth: Overall Inequality

| | Original Eq | Counterfactual | Change |
|---|---|---|---|
| **Panel A. Wage** | | | |
| Gini Index | 0.432 | 0.223 | -48.4% |
| P90/P10 | 7.293 | 2.647 | -63.7% |
| P90/P50 | 4.521 | 1.901 | -58.0% |
| **Panel B. Real Income** | | | |
| Gini Index | 0.420 | 0.207 | -50.7% |
| P90/P10 | 7.000 | 2.452 | -65.0% |
| P90/P50 | 4.339 | 1.788 | -58.8% |

# Counterfactual 1: Patents Reduce to 2005 Level

Table 18: Eliminating Innovation Growth: Gini Coefficient by Hukou Type

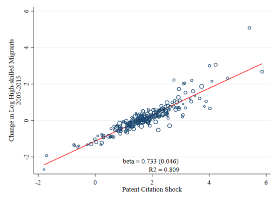
| | Original Eq | Counterfactual | Change |
|---|---|---|---|
| **Panel A. Agr Hukou** | | | |
| Wage | 0.381 | 0.159 | -58.3% |
| Real Income | 0.368 | 0.156 | -57.6% |
| Welfare | 0.068 | 0.074 | 8.8% |
| **Panel B. Non-agr Hukou** | | | |
| Wage | 0.165 | 0.261 | 58.1% |
| Real Income | 0.160 | 0.258 | 61.3% |
| Welfare | 0.102 | 0.185 | 81.4% |

# Counterfactual 2: Channel Decomposition

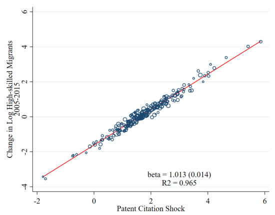
- There are three main channels for technology to affect migration
  - Wage effect attracts migration
  - Direct amenity effect attracts migration
  - Indirect amenity effect through changes of skill ratio
- We now try to decompose the overall effect of tech growth on migration into these three channels

# Counterfactual 2: Channel Decomposition

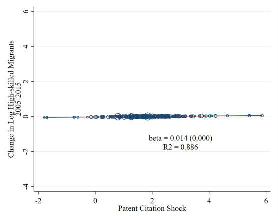
[Original Counterfactual]




[Fixed skill ratio]




[Fixed skill ratio, a]




[Fixed skill ratio, w, a]

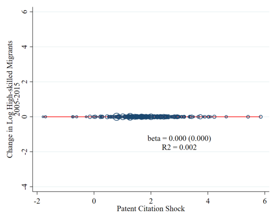
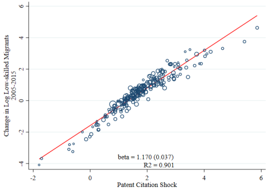


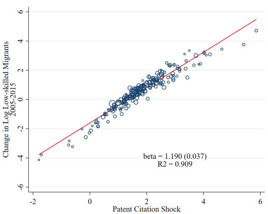

Figure 7: Channel Analysis of Patent's Effect on High-skilled Migration

# Counterfactual 2: Channel Decomposition

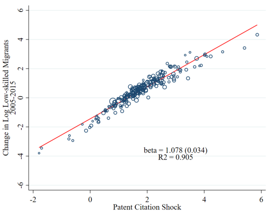
[Original Counterfactual]




[Fixed skill ratio]




[Fixed skill ratio, a]




[Fixed skill ratio, w, a]

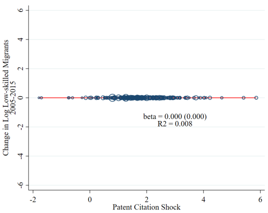



Figure 8: Channel Analysis of Patent's Effect on Low-skilled Migration

# Counterfactual 2: Channel Decomposition

Table 19: Channel Decomposition of Patent's Effect on Migration

| | Indirect Amenity Effect | Direct Amenity Effect | Wage Effect |
|---|---|---|---|
| **Panel.A High-skilled Migration** | | | |
| Coefficient Change | -0.293 | 0.997 | 0.029 |
| Proportional Explained | -40.0% | 136.0% | 4.0% |
| **Panel.B Low-skilled Migration** | | | |
| Coefficient Change | -0.016 | 0.057 | 1.129 |
| Proportional Explained | -1.4% | 4.9% | 96.5% |

# Conclusion

- Patent growth in China during 2005 to 2015 increased wages for both low- and high-skilled workers
- Low-skilled workers care more about wages while high-skilled workers care more about amenities
- Patent growth attracted more low-skilled workers, reduced the skill ratio, which reduced amenities and discourages high-skilled migrants
- Technology growth in China during the last twenty years DID NOT lead to a gentrification and benefited workers with different skills

# Summary Statistics

| VARIABLES | (1) N | (2) Mean | (3) SD | (4) Min | (5) Max |
|---|---|---|---|---|---|
| Migrant/Employment | 609 | 0.16 | 0.16 | 0.00 | 0.90 |
| High-Skilled Migrant/High-Skilled Employment | 609 | 0.08 | 0.09 | 0.00 | 0.59 |
| Low-Skilled Migrant/Low-Skilled Employment | 609 | 0.19 | 0.19 | 0.00 | 0.96 |
| Citations of Patents | 571 | 2903.20 | 11265.63 | 2.00 | 157306.00 |
| High-skilled wage | 595 | 48127.27 | 14217.58 | 15928.67 | 122615.09 |
| Low-skilled wage | 595 | 39474.81 | 11309.51 | 6007.17 | 91138.81 |
| City-level average house price | 570 | 4822.948 | 3086.891 | 1589.353 | 33942.34 |
| Doctors per 10,000 residents | 576 | 20.37 | 8.18 | 6.92 | 75.19 |
| Hospitals per 10,000 residents | 576 | 0.60 | 0.65 | 0.09 | 6.89 |
| Kilometers of road per 10,000 residents | 575 | 33.38 | 18.72 | 1.44 | 152.09 |
| Highway passengers per 10,000 residents | 574 | 24.30 | 121.30 | 1.15 | 2855.72 |
| High-speed Railway Connected | 577 | 0.41 | 0.49 | 0.00 | 1.00 |
| PM 2.5 | 572 | 44.64 | 20.05 | 4.15 | 101.19 |
| Heavily polluted days | 576 | 6.85 | 10.91 | 0.00 | 55.89 |
| Polluted days | 576 | 70.37 | 56.32 | 0.00 | 237.05 |
| Teacher-student ratio in primary schools | 577 | 0.06 | 0.01 | 0.00 | 0.13 |
| Teacher-student ratio in middle schools | 576 | 0.08 | 0.02 | 0.00 | 0.20 |
| Number of colleges | 565 | 8.45 | 14.55 | 1.00 | 90.00 |
| Number of Project 985 universities | 578 | 0.13 | 0.64 | 0.00 | 8.00 |
| Number of Project 211 universities | 578 | 0.38 | 1.75 | 0.00 | 23.00 |
| Log Slope | 575 | 0.94 | 1.13 | -6.89 | 2.91 |

# Initial Skilled Ratio and Change in Skilled Ratio


Figure 9: Initial Skilled Ratio and Change in Skilled Ratio

# Model Fit

Table 20: Model Fit

| Variables | Model | Data | Difference |
|---|---|---|---|
| Total Migrants | 27410758 | 26910338 | 1.86% |
| Total High-skill Migrants | 5956344 | 5859850 | 1.65% |
| Total Low-skill Migrants | 21454414 | 21050488 | 1.92% |
| Total Migration in East | 18515916 | 18109484 | 2.24% |
| Total Migration in Middle | 4098104 | 4014531 | 2.08% |
| Total Migration in Northeast | 1081882 | 1065493 | 1.54% |
| Total Migration in West | 3714857 | 3720829 | 0.16% |
| Mean Wages of High-skill in Agr | 12723 | 12727 | -0.036% |
| Mean Wages of High-skill in Non-agr | 56917 | 56829 | 0.155% |
| Mean Wages of Low-skill in Agr | 12723 | 12727 | -0.036% |
| Mean Wages of Low-skill in Non-agr | 46861 | 46764 | 0.21% |
| Mean Housing Price | 6161.8 | 6161.4 | -0.0067% |
| Mean Amenity | 1.352 | 1.351 | 0.07% |

Back

# Model Fit

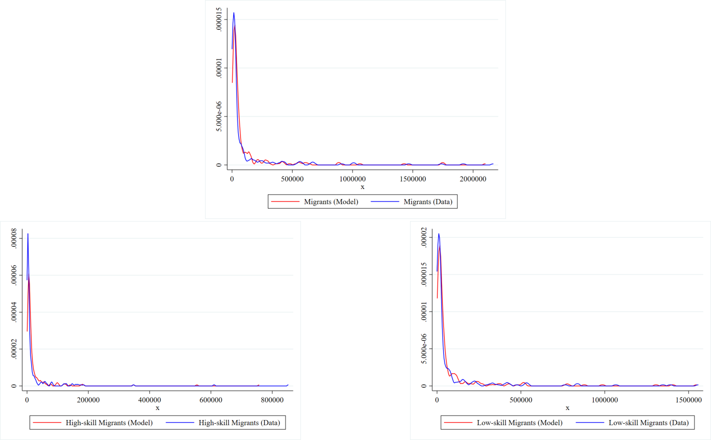



Figure 10: Model Fit of Migrants

Back

# Model Fit

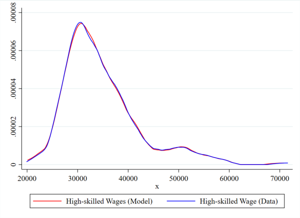
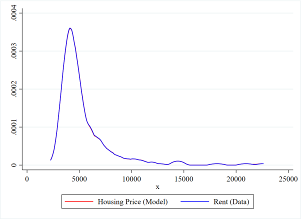




Figure 11: Model Fit of Wages

Back

# Model Fit



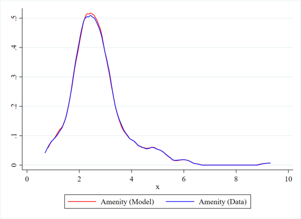




Figure 12: Model Fit of Housing Price and Amenity

▸ Back

# Spatial Distribution of Patent Shock

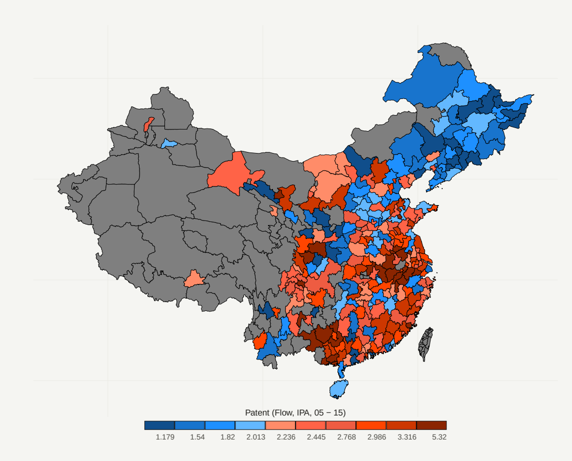



Figure 13: Spatial Distribution of Log Patent Growth (2005 - 2015)

Back

# Descriptive Analysis

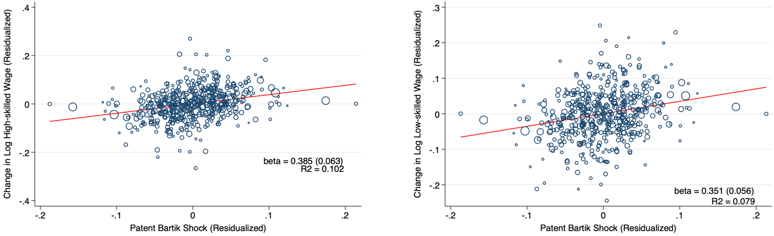



Figure 14: Effect of Patent Shock on Wages for High- and Low-skilled Workers

Back

# Descriptive Analysis

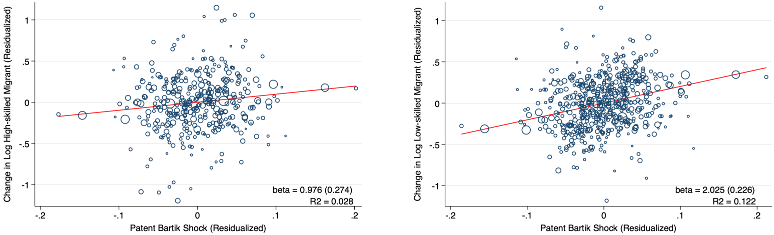



Figure 15: Effect of Patent Shock on Number of High- and Low-skilled Migrants

▸ Back

# Descriptive Analysis

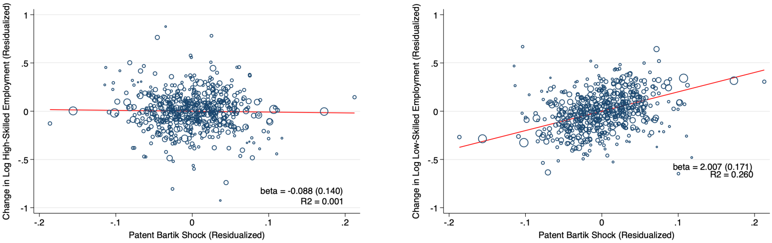



Figure 16: Effect of Patent Shock on Number of High- and Low-skilled Employment

Back

# Descriptive Statistics

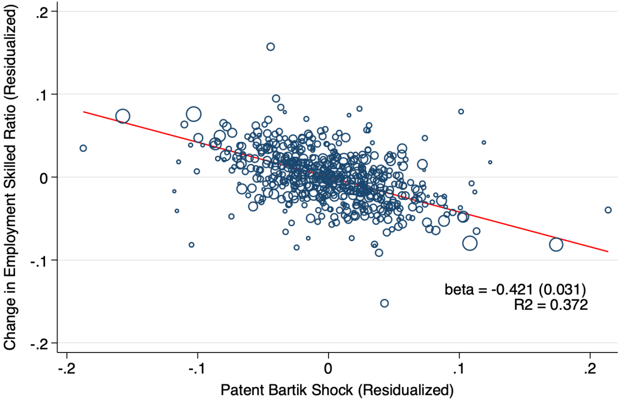



Figure 17: Patent Shock and Change in Skilled Ratio

Back

# Descriptive Analysis

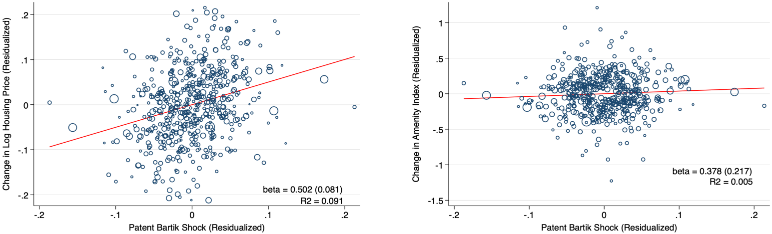



Figure 18: Effect of Patent Shock on Housing Price and Amenity

Back

# Patent growth and labor supply

| VARIABLES | (1) Δ Log Employment | (2) Δ Log High-Skilled Employment | (3) Δ Log Low-Skilled Employment | (4) Δ Log High-skilled Migrants | (5) Δ Log Low-skilled Migrants | (6) Δ Employment Skilled Ratio |
|---|---|---|---|---|---|---|
| *Panel A: OLS* | | | | | | |
| Δ Log(Patent) | 0.0903** | 0.0157 | 0.127** | 0.114 | 0.192*** | -0.0267*** |
| | (0.0397) | (0.0362) | (0.0509) | (0.0817) | (0.0724) | (0.00993) |
| *Panel B: Reduced Form* | | | | | | |
| Patent shock | 1.256*** | -0.0877 | 2.007*** | 0.976** | 2.025*** | -0.421*** |
| | (0.235) | (0.199) | (0.242) | (0.389) | (0.320) | (0.0435) |
| *Panel C: IV* | | | | | | |
| Δ Log(Patent) | 1.103*** | -0.111 | 1.795*** | 0.675** | 1.874*** | -0.373*** |
| | (0.302) | (0.187) | (0.442) | (0.312) | (0.540) | (0.0888) |
| Year FE | X | X | X | X | X | X |
| City FE | X | X | X | X | X | X |

Back

# Patent growth, wage, housing price, and amenity

| VARIABLES | (1) Δ Log High-skilled Wage | (2) Δ Log Low-skilled Wage | (3) Δ Log(Housing Price) | (4) Δ Amenity Index |
|---|---|---|---|---|
| *Panel A: OLS* | | | | |
| Δ Log(Patent) | -0.00123 | -0.00769 | -0.0249 | 0.0954 |
| | (0.0162) | (0.0174) | (0.0198) | (0.0799) |
| *Panel B: Reduced Form* | | | | |
| Patent shock | 0.385*** | 0.351*** | 0.502*** | 0.378 |
| | (0.0887) | (0.0795) | (0.115) | (0.307) |
| *Panel C: IV* | | | | |
| Δ Log(Patent) | 0.402*** | 0.348*** | 0.469** | 0.428 |
| | (0.146) | (0.134) | (0.199) | (0.306) |
| Year FE | X | X | X | X |
| City FE | X | X | X | X |

Back

# Estimation of the Labor Demand in the Non-agricultural Sector (Patent)

Table 21: Estimation of the Labor Demand in the Non-agricultural Sector

| VARIABLES | (1) Δ Log High-skilled Wage | (2) Δ Log Low-skilled Wage | (3) Δ Log High-skilled City Population | (4) Δ Log Low-skilled City Population | (5) Δ Log Patent | (6) Δ Log High-skilled Wage | (7) Δ Log Low-skilled Wage |
|---|---|---|---|---|---|---|---|
| Δ Log Patent (Flow, IPA) | 0.114*** | 0.114*** | | | | 0.514*** | 0.481*** |
| | (0.0116) | (0.0132) | | | | (0.0877) | (0.0795) |
| Δ Log High-skilled City Population | 0.00306 | 0.00559 | | | | -0.777* | -0.556 |
| | (0.0224) | (0.0254) | | | | (0.435) | (0.394) |
| Δ Log Low-skilled City Population | -0.0584** | -0.0635** | | | | 0.107 | -0.0860 |
| | (0.0282) | (0.0319) | | | | (0.426) | (0.387) |
| Patent shock | | | -0.0207 | 0.100 | 0.978*** | | |
| | | | (0.125) | (0.0989) | (0.175) | | |
| Migrant bartik for high-skill workers | | | 0.564*** | 0.404*** | 0.659*** | | |
| | | | (0.142) | (0.112) | (0.199) | | |
| Migrant bartik for low-skill workers | | | -0.162 | 0.721** | -0.253 | | |
| | | | (0.357) | (0.282) | (0.500) | | |
| Constant | 0.543*** | 0.545*** | -0.226 | -0.832*** | -0.371 | 0.185* | 0.174* |
| | (0.0233) | (0.0263) | (0.348) | (0.275) | (0.488) | (0.0980) | (0.0889) |
| Observations | 484 | 484 | 484 | 484 | 484 | 484 | 484 |
| R-squared | 0.174 | 0.142 | | | | | |
| Model | OLS | OLS | First stage | First stage | First stage | IV GMM | IV GMM |
| Cragg-Donald Wald F | | | | | | 2.266 | 2.266 |
| Sanderson-Windmeijer F | | | 6.819 | 9.238 | 9.247 | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Back

# Estimation of the Amenity Market (Patent)

Table 22: Estimation of the Amenity Market

| VARIABLES | (1)<br>Δ Amenity Index | (2)<br>Δ Log High-skilled City Population Ratio | (3)<br>Δ Log Patent | (4)<br>Δ Amenity Index |
|---|---|---|---|---|
| Δ Log Patent | 0.266*** | | | 0.580*** |
| | (0.0360) | | | (0.0960) |
| Δ High-skilled City Population Ratio | 0.0274 | | | 1.596 |
| | (0.383) | | | (1.872) |
| Patent shock | | -0.0812*** | 0.723*** | |
| | | (0.0236) | (0.220) | |
| Wage Bartik IV for high-skilled workers | | 0.0872 | -0.523 | |
| | | (0.0577) | (0.536) | |
| Wage Bartik IV for low-skilled workers | | 0.145** | 1.529** | |
| | | (0.0724) | (0.672) | |
| Constant | 0.325*** | -0.0283** | 0.00193 | -0.312** |
| | (0.0726) | (0.0140) | (0.130) | (0.132) |
| Observations | 484 | 484 | 484 | 484 |
| R-squared | 0.104 | | | |
| Model | OLS | First stage | First stage | IV GMM |
| Cragg-Donald Wald F | | | | 7.948 |
| Sanderson-Windmeijer F | | 12.05 | 19.40 | |

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Back

# Tests of Bartik Shocks

Table 23: Summary Statistics of Bartik Shock

| Bartik | (1)<br>Citation | (2)<br>Migrant<br>High-Skilled | (3)<br>Migrant<br>Low-Skilled | (4)<br>Wage<br>High-Skilled | (5)<br>Wage<br>Low-Skilled |
|---|---|---|---|---|---|
| Panel A: 2010 | | | | | |
| Mean | 1.269 | 0.727 | 0.729 | 0.669 | 0.665 |
| Standard deviation | 0.326 | 0.269 | 0.277 | 0.102 | 0.102 |
| Effective Sample Size (1/HHI) | 13.834 | 7.168 | 4.295 | 10.131 | 5.763 |
| Panel B: 2015 | | | | | |
| Mean | 1.481 | 0.741 | 0.724 | 1.143 | 1.136 |
| Standard deviation | 0.533 | 0.434 | 0.447 | 0.099 | 0.097 |
| Effective Sample Size (1/HHI) | 19.058 | 8.999 | 5.228 | 10.597 | 6.046 |

Back

# Tests of Bartik Shocks

Table 24: Balance Test of Bartik-type Shock of Citation

| VARIABLES | 2010 | | 2015 | |
|---|---|---|---|---|
| | Coefficient | S.E. | Coefficient | S.E. |
| GDP Growth Rate | 0.00150 | (0.00303) | 0.000899 | (0.00244) |
| Log Population Density | 0.00166 | (0.00871) | -0.00254 | (0.00493) |
| Log Fiscal Expenditure | 0.0168 | (0.0226) | -0.00361 | (0.0283) |
| Log Retail Sale | -0.00131 | (0.0128) | -0.0140 | (0.00871) |
| Log Above-Scale Firm Profit | -0.0956 | (0.0868) | 0.0287 | (0.108) |
| Log Number of College Students | -0.00429 | (0.0204) | -0.00376 | (0.0200) |

Back

# Tests of Bartik Shocks

Table 25: Balance Test of Bartik-type Shock of Migrants

| VARIABLES | 2010 Coefficient | S.E. | 2015 Coefficient | S.E. |
|---|---|---|---|---|
| *Panel A: High-Skilled* | | | | |
| GDP Growth Rate | -0.00278 | (0.00673) | 0.00250 | (0.00520) |
| Log Population Density | -0.00762 | (0.00779) | -0.0202* | (0.0108) |
| Log Fiscal Expenditure | -0.0415 | (0.0573) | -0.0165 | (0.0942) |
| Log Retail Sale | 0.0195 | (0.0496) | -0.0145 | (0.0665) |
| Log Above-Scale Firm Profit | -0.0109 | (0.118) | 0.0746 | (0.106) |
| Log Number of College Students | 0.725 | (0.476) | 0.258 | (0.559) |
| *Panel B: Low-Skilled* | | | | |
| GDP Growth Rate | -0.00629** | (0.00230) | -0.00404 | (0.00298) |
| Log Population Density | 0.00242 | (0.00196) | 0.00609 | (0.00486) |
| Log Fiscal Expenditure | -0.0382 | (0.0249) | -0.0357 | (0.0395) |
| Log Retail Sale | -0.00912 | (0.0123) | -0.0432* | (0.0232) |
| Log Above-Scale Firm Profit | -0.0790 | (0.0683) | 0.0483 | (0.138) |
| Log Number of College Students | 0.403 | (0.245) | 0.143 | (0.261) |

Back

# Tests of Bartik Shocks

Table 26: Balance Test of Bartik-type Shock of Wages

| VARIABLES | 2010 | | 2015 | |
|---|---|---|---|---|
| | Coefficient | S.E. | Coefficient | S.E. |
| *Panel A: High-Skilled* | | | | |
| GDP Growth Rate | -0.00591 | (0.00836) | -0.00479 | (0.00536) |
| Log Population Density | 0.00448 | (0.0120) | -0.0149 | (0.0133) |
| Log Fiscal Expenditure | 0.0220 | (0.0703) | -0.0140 | (0.117) |
| Log Retail Sale | -0.00487 | (0.0299) | -0.0370 | (0.0660) |
| Log Above-Scale Firm Profit | -0.114 | (0.273) | 0.144 | (0.289) |
| Log Number of College Students | -1.807** | (0.683) | -2.081*** | (0.688) |
| *Panel B: Low-Skilled* | | | | |
| GDP Growth Rate | 0.0124 | (0.0110) | 0.00377 | (0.00721) |
| Log Population Density | -0.00680 | (0.00533) | -0.0121* | (0.00609) |
| Log Fiscal Expenditure | -0.00151 | (0.0765) | -0.0493 | (0.113) |
| Log Retail Sale | -0.0129 | (0.0363) | -0.0468 | (0.0525) |
| Log Above-Scale Firm Profit | -0.212 | (0.220) | 0.102 | (0.377) |
| Log Number of College Students | -0.672* | (0.327) | -0.653 | (0.444) |

Back

# References

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4):588–638.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4):841–890.

Card, David. 2009. "Immigration and Inequality." *American Economic Review* 99 (2):1–21.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review* 106 (3):479–524.