Migration Restrictions and the Migrant-Native Wage Gap: The Role of Wage Setting and Sorting

Naijia Guo[1] Li Zhang[2] Rongjie Zhang[3] Ben Zou[4]

[1]HKU [2]UCL [3]Tsinghua [4]Purdue

August, 2024

Migration Restrictions and Wage Gap

- Large-scale migration is a salient feature of the economy
  - One in seven individuals in the world is a migrant. (Bell and Charles-Edwards, 2013)
  - Around 56% of residents in China's largest cities were born somewhere else.

Migration Restrictions and Wage Gap

- Large-scale migration is a salient feature of the economy
  - One in seven individuals in the world is a migrant. (Bell and Charles-Edwards, 2013)
  - Around 56% of residents in China's largest cities were born somewhere else.
- Overwhelming evidence of migrant-native wage gap. (Clarke et al., 2019; Cupak et al., 2023; Dostie et al., 2023; Ma, 2018; Zhang et al., 2016; Zhu, 2016.)

Migration Restrictions and Wage Gap

- Large-scale migration is a salient feature of the economy
  - One in seven individuals in the world is a migrant. (Bell and Charles-Edwards, 2013)
  - Around 56% of residents in China's largest cities were born somewhere else.
- Overwhelming evidence of migrant-native wage gap. (Clarke et al., 2019; Cupak et al., 2023; Dostie et al., 2023; Ma, 2018; Zhang et al., 2016; Zhu, 2016.)
- Discriminatory policies restricting migration at destination labor markets.
  - e.g., H1B visa in the U.S., guest worker programs in many Gulf countries.
  - In China, *hukou* policies restrict *internal* migration.

Research Questions

- What factors drive the migrant-native wage gap?
- How do internal migration restrictions affect the migrant-native wage gap?

Migrant-native Wage Gap in Urban China

- Existing studies focus on differences in worker characteristics and residual wage gap. (Chen and Zhang, 2018; Lee, 2012; Ma, 2018; Song 2016; Zhang et al., 2016)
- This paper: employer-employee interactions
  - Leverage the first employer-employee matched panel in China.
  - Recover unobserved worker skills.
  - Wage setting: Same employer pays differential wage premiums to migrants and natives with the same skill.
  - Sorting: Migrants and natives work for employers that pay different premiums.
- Reasons for wage setting and sorting
  - Exploit a quasi-natural experiment: *hukou* quota tightening

Setting and Data

- Employer-employee linked panel data in a large Chinese city
  - The *formal* sector covered by the Housing Provident Fund.
  - Positively selected, but a relevant population with a pathway to a local *hukou*.
- An employer-sponsored *hukou* quota system
  - Allocated by the government, tilted towards the public sector.
  - A major way to obtain *hukou* besides family-based *hukou* acquisition
  - A policy change that further tightens and tilts the quotas.

1: Components of the Migrant-Native Wage Gap

- AKM model + Oaxaca-Blinder decomposition
  - ln wage = Worker's Skill (observed & unobserved) + Employer premium + ε.
  - Gap in employer premium = Gap in within-employer wage setting + Gap in sorting into different employers.
  - Sorting: Skill-based + residual.
- Average wage gap: Migrants - native = +19pp
  - Migrants have higher skills (+21pp).
  - Skill-based sorting (+6pp).
  - Wage penalty for migrants (-8pp).
    - Wage setting: -3pp.
    - Residual sorting: -5pp.

2: Causal Impacts of *Hukou* Quota Tightening

- Policy change that reduced 40% of *hukou* quota, remaining quota further tilting towards the public sector.
  - Wages increase for migrants relative to natives by 5.6%
  - Wage setting increases by 3.7pp, and the increase is larger
    - in the private sector
    - for skilled and young migrants
  - Residual sorting declines by 1.4pp
    - Young and skilled migrants increasingly work in the public sector with more *hukou* quota but lower productivity

Literature and Contributions

- Migrant-native wage gap
  - Previous studies focus on the effects of worker characteristics, network, skill mismatch, and search friction (Bartolucci, 2014; Hirsch and Jahn, 2012; Ma, 2018; Pendakur and Woodcock, 2010; Picot and Piraino, 2013; Zhang et al., 2016; Zhu, 2016)
  - **Contribution**: We leverage the employer-employee linked data to analyze the role of wage setting and sorting effects
- AKM decomposition for group wage gap (gender gap: Card et al., 2013, 2016; ethnicity gap: Gerard et al., 2021)
  - **Contribution**: Exploit a policy change to analyze the sources of wage setting and sorting effects.
- Migration restriction and its implications
  - Employer-sponsored work permits: H1B visa in the U.S.(Khanna et al., 2022; Pei, 2024); Guest worker programs in Gulf countries (Naidu et al., 2016).
  - Hukou system in China, mostly focused on rural-urban (low-skilled) migrants (Meng 2012, Fan 2019, Gai et al. 2021, Sieg et al. 2023)
  - **Contribution**: How hukou quota affects wages and sorting.

Data

- Matched employer-employee panel 2006-2014 in one major city
  - Employer ID, worker ID, Housing Provision Fund contribution
    - HPF is part of the employment-based social insurance, and its contribution is a fixed % of labor earnings (salary+bonuses).
  - Worker information: migration status (from national ID), age, gender.
  - Employers: sector and industry for most. Include firms, not-for-profits, and government agencies.
- Sample: 22-50 years old.
  - Drop labor dispatch firms and human resources service firms.
- Definition of migrant
  - Those who were born in other cities, regardless of current hukou status.
  - Most are high-skilled workers.

Summary Statistics (2010)

Panel A: Worker	Migrants	Natives
Number of individuals	1,155,873	1,658,170
Percent of females (%)	44.0	46.9
Age	30.8	36.0
	(6.0)	(8.5)
Wage (CNY)	72,086	62,424
	(64,421)	(53,200)
Percent of public sector (%)	37.5	68.9
	(48.4)	(46.3)
Percent of job switches in the past year (%)	8.9	7.1
	(28.4)	(25.7)
Percent of leaving the sample next year (%)	9.6	6.9
	(29.5)	(25.4)
Panel B: Employer	**Public**	**Private**
Number of employers	11,654	17,429
Percent of enterprises (%)	45.0	99.0
Employer size	114.9	59.2
	(337.4)	(181.6)
Average wage (CNY)	58,996.3	59,867.0
	(34,474.1)	(44,049.8)
Proportion of female workers (%)	51.5	45.9
	(21.9)	(23.5)
Proportion of native workers (%)	74.9	42.4
	(25.2)	(29.4)
Average age	36.7	32.4
	(4.5)	(4.3)
Turnover rate (%)	27.4	66.2
	(227.1)	(614.2)

Supplemental OLS results | Compared with aggregate data | Sectoral difference in wage gap

Outline of the Presentation

The AKM Model

$$\ln y_{ijt} = \alpha_i + X'_{it}\beta + \psi_j + \varepsilon_{it}$$

- $\ln y_{ijt}$: log annual wage of individual i in employer j in year t.
- α_i: person effect (PE), capturing time-invariant skills of i.
- X_{it}: time-varying observable chars, incl. polynomials of age and their interactions with gender, as well as the year dummies and their interactions with gender.
- $\alpha_i + X'_{it}\beta$: person skill.
- ψ_j: employer effect (EE), capturing wage premiums.

Assumptions and validations Decomposition of wage variance

Group-Specific AKM

$$\ln y_{gijt} = \alpha_{gi} + X'_{git}\beta_g + \psi_{g_j} + \varepsilon_{git}$$

- Separately estimate for two groups: Migrants and Natives
- To make PE and EE from separate groups comparable:
  - We need benchmark employers with zero wage premium.
  - Set EE=0 for private employers with average wage btw. 5-10th percentiles, i.e., low-wage employers have no wage premium.

Procedure of normalization | Normalized EE and Wage

Oaxaca-Blinder Decomposition of Group Wage Gap

$$
\begin{aligned}
& E\left[\ln y_{Mijt}\right]-E\left[\ln y_{Nijt}\right] \\
& =\underbrace{\alpha_M-\alpha_N}_{\text{gap in person effect}}+\underbrace{\bar{X}_M^{\prime} \beta_M-\bar{X}_N^{\prime} \beta_N}_{\text{gap in covariates}}+\underbrace{\sum_j \psi_j^M \pi_{Mj}-\sum_j \psi_j^N \pi_{Nj}}_{\text{gap in employer premium}} \\
& =\underbrace{\alpha_M-\alpha_N}_{\text{gap in person effect}}+\underbrace{\bar{X}_M^{\prime} \beta_M-\bar{X}_N^{\prime} \beta_N}_{\text{gap in covariates}}+\underbrace{\sum_j\left(\psi_j^M-\psi_j^N\right) \pi_{Mj}}_{\text{wage setting}}+\underbrace{\sum_j \psi_j^N\left(\pi_{Mj}-\pi_{Nj}\right)}_{\text{sorting}}
\end{aligned}
$$

- $\pi_{gj}=\frac{N_{gj}}{N_g}$: employer j's share of group g workers.
- Wage setting: Within the same employer, same-skilled workers in certain groups are paid less.
- Sorting: Workers in certain groups are less likely to be hired by high-premium employers.

Sorting

$$\sum_j \psi_j^N \cdot (\pi_{Mj} - \pi_{Nj}) = \underbrace{\sum_j \psi_j^N (\pi^*_{Mj} - \pi^*_{Nj})}_{\text{Skill-based sorting}} + \underbrace{\sum_j \psi_j^N ((\pi_{Mj} - \pi^*_{Mj}) - (\pi_{Nj} - \pi^*_{Nj}))}_{\text{Residual sorting}}$$

- π^*_{gj}: counterfactual share of group g in employer j.
  - Employer would achieve the same skill mix of workers,
  - but select workers regardless of their migration status.
  - $\pi^*_{gj} \equiv \frac{\sum_z N_{zj} \pi_{gz}}{\sum_z N_{zj}}$, where N_{zj} is the number of workers in skill group z in employer j, π_g is the share of group g worker of skill z in the labor market.
- Skill-based sorting: higher-premium employers tend to have a greater demand for skilled workers.
- Residual sorting: employers' discriminatory hiring policies + workers' heterogeneous preferences.

Migrant-Native Wage Gap Decomposition

	Baseline	Alternative Decomposition
Log Wage Gap (Migrant - Native, same below)	0.194***	0.194***
	[0.002]	[0.002]
1. Person Effect	0.326***	0.326***
	[0.016]	[0.016]
2. Covariates	-0.120***	-0.120***
	[0.006]	[0.006]
3. Employer Effect	-0.012	-0.012
	[0.015]	[0.015]
3.1. Wage Setting Effect	-0.026*	-0.039***
	[0.015]	[0.015]
3.2. Sorting	0.014***	0.027***
	[0.003]	[0.005]
3.2.1. Skill-Based Sorting	0.064***	0.069***
	[0.002]	[0.002]
3.2.2. Residual Sorting	-0.051***	-0.042***
	[0.003]	[0.003]

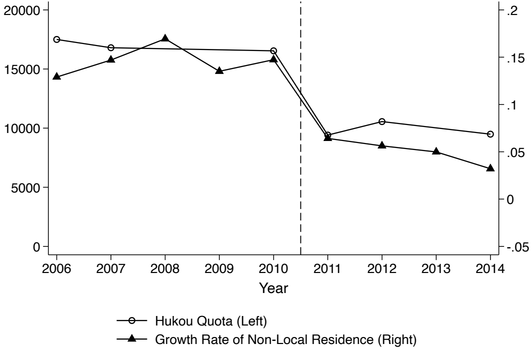
- Positive wage gap arises from higher person effects of migrants.
- Wage setting & residual sorting are negative.

Hukou System in the Studied City

- *Hukou* policy in the studied city
  - Local *hukou* grants access to local public goods, highly valuable for migrants.
  - Employment-based pathway mostly operates through a quota system.
    - Quotas for fresh college graduates, talents programs, quasi-*hukou* permit, etc.
    - Government allocates quota, favors the public sector.
    - Quota typically explicit in job posting.
  - *Hukou* is portable, but usually, workers are locked in for 3-5 years.
  - Young and skilled migrants are more affected.
    - Quota often reserved for the young and the skilled.
    - Young and skilled have a higher WTP for *hukou*.
- *Hukou* quota tightening in 2011
  - Mainly affect the private sector.

Tightening of *Hukou* Quota

- By the end of 2009, the population in the studied city had already reached the target that had been set for 2020.
- In 2011, the government tightened *hukou* quotas by 40% to reduce migration inflow.
  - Remaining quota is further tilted towards the public sector.



Source: Municipal Statistical Yearbooks.
City-wide population: 2000-10 growth: 45%, 2010-20 growth: 11% Labor Supply Migrant with *hukou*

Worker's Labor Supply Problem

For worker i in type g (e.g., migrants vs. natives, high-skilled vs. low-skilled, young vs. old), the indirect utility of working at firm j is

$$u_{igj} = \beta_g \ln(w_{gj}) + \gamma_g a_{gj} + \epsilon_{igj}$$

where a_{gj} is the hukou quota offered by firm j. β_g and γ_g capture the type-specific preferences for wage and hukou quota, respectively.

Workers have logit choice probabilities as follows:

$$\begin{aligned} p_{gj} &= \frac{\exp(\beta_g \ln(w_{gj}) + \gamma_g a_{gj})}{\sum_{k=1}^{J} \exp(\beta_g \ln(w_{gk}) + \gamma_g a_{gk})} \\ &\approx \lambda_g \exp(\beta_g \ln(w_{gj}) + \gamma_g a_{gj}) \end{aligned}$$

Firm's Optimization Problem

Firms have the linear production function

$$Y_j = T_j(\sum_g z_g N_{gj})$$

where T_j is the labor productivity of firm j, z_g represents the efficient units of labor for type g worker, and N_{gj} is the labor supply for type g worker in firm j.

Assuming that the number of hukou quotas a_{gj} and its associated cost c are exogenous, firm's problem is to post a set of type-specific wages that minimize the cost of labor

$$\min_{w_{gj}} \sum_g (w_{gj}(1 + a_{gj}c))N_{gj}(w_{gj}) \quad \text{such that} \quad T_j(\sum_g z_g N_{gj}) \geq Y$$

Equilibrium Outcomes

In the equilibrium, the optimal wage satisfies

$$\ln w_{gj} = \ln\left(\frac{\beta_g}{1+\beta_g}\right) + \ln(z_g) + \ln(T_j) - \ln(1 + a_{gj}c)$$

The labor supply of type g worker in firm j is:

$$\begin{aligned}\ln N_{gj}(w_{gj}) = \beta_g & \left(\ln\left(\frac{\beta_g}{1+\beta_g}\right) + \ln(z_g) + \ln(T_j) - \ln(1 + a_{gj}c)\right) \\ & + \gamma_g a_{gj} + \ln(N_g \lambda_g)\end{aligned}$$

Model Predictions

Cross-sectional predictions

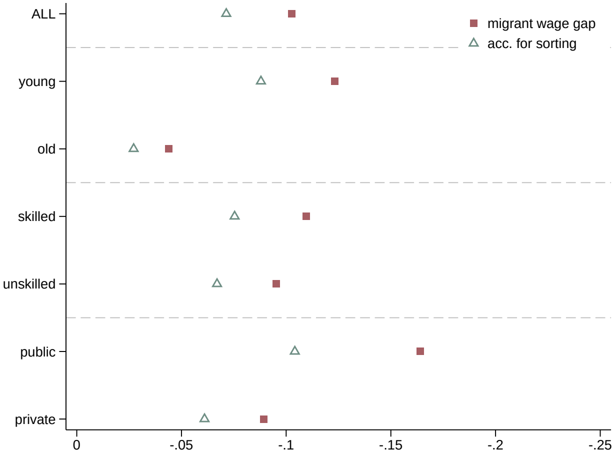
- Since firms provide hukou to migrants but not natives, migrants will be paid lower wages than natives in the same firm – negative wage setting effect.
- The negative wage-setting effect is particularly strong for young and skilled workers and those in the public sector, as hukou quotas are typically tilted towards these groups.
- If firms with higher employer premiums have fewer hukou quotas (e.g., private firms), migrants are more likely to sort into low premium firms than natives – negative residual sorting effect.
- Suppose young and skilled workers have a higher WTP for hukou quota (higher γ_g/β_g), the magnitude of residual sorting will be larger for them.

Descriptive Evidence: Migrant-native Wage Gap

$$\ln y_i = \gamma \cdot \mathsf{Mig}_i + \beta_X X_i + \beta_{Id} \hat{\alpha}_i^d \quad (+\zeta_j) + \varepsilon_i,$$

- Controlling for deciles of person effect ($\hat{\alpha}_i^d$) controls for workers' unobserved characteristics and skilled-based sorting.
- γ_1 captures the wage gap when employer FE (ζ_j) is not included (wage-setting + residual sorting), and γ_2 captures the gap when employer FE is included (wage-setting).
- Predictions
  - γ_2 is negative.
  - γ_1 is larger in magnitude than γ_2.
  - γ_2 are larger in magnitude for the young and skilled workers working in the public sector
  - $\gamma_2 - \gamma_1$ are also larger for these workers.

Figure: Descriptive Evidence of Wage Setting and Residual Sorting



Model Predictions

Time-varying predictions related to hukou quota change

- The reduction in hukou quota will increase the wage-setting effect, particularly for private firms and young, skilled workers.
- If the reduction in hukou quota primarily targets the private sector (firms with higher employer premiums), residual sorting will decline, particularly for young and skilled workers.

Changes in Wage Gap Components

	(1) Pre-shock	(2) Post-shock	(3) Difference (Post-Pre)
Wage Gap (Migrant - Native, same below)	0.194***	0.205***	0.011***
	[0.002]	[0.002]	[0.002]
1. Person Effect Gap	0.326***	0.277***	-0.049**
	[0.016]	[0.014]	[0.021]
2. Covariates Gap	-0.120***	-0.069***	0.051***
	[0.006]	[0.003]	[0.007]
3. Employer Effect Gap	-0.012	-0.003	0.009
	[0.015]	[0.013]	[0.020]
3.1. Wage Setting Effect	-0.026*	0.011	0.037*
	[0.015]	[0.013]	[0.020]
3.2. Sorting Effect	0.014***	-0.014***	-0.028***
	[0.003]	[0.002]	[0.004]
3.2.1. Skill-Based Sorting Effect	0.064***	0.051***	-0.014***
	[0.002]	[0.002]	[0.003]
3.2.2. Residual Sorting Effect	-0.051***	-0.065***	-0.014***
	[0.003]	[0.002]	[0.003]

- Wage setting: migrants' wages ↑ by 3.7pp relative to natives.
  - Compensating for fewer *hukou* quotas
  - Monopsony power ↓ for employers that have quota reduced

Changes in Wage Gap Components

	(1) Pre-shock	(2) Post-shock	(3) Difference (Post-Pre)
Wage Gap (Migrant - Native, same below)	0.194***	0.205***	0.011***
	[0.002]	[0.002]	[0.002]
1. Person Effect Gap	0.326***	0.277***	-0.049**
	[0.016]	[0.014]	[0.021]
2. Covariates Gap	-0.120***	-0.069***	0.051***
	[0.006]	[0.003]	[0.007]
3. Employer Effect Gap	-0.012	-0.003	0.009
	[0.015]	[0.013]	[0.020]
3.1. Wage Setting Effect	-0.026*	0.011	0.037*
	[0.015]	[0.013]	[0.020]
3.2. Sorting Effect	0.014***	-0.014***	-0.028***
	[0.003]	[0.002]	[0.004]
3.2.1. Skill-Based Sorting Effect	0.064***	0.051***	-0.014***
	[0.002]	[0.002]	[0.003]
3.2.2. Residual Sorting Effect	-0.051***	-0.065***	-0.014***
	[0.003]	[0.002]	[0.003]

- Gap in residual sorting decreased by 1.4pp.
  - Remaining hukou quotas are concentrated in the public sector, which has relatively low productivity.
  - Migrants become even more unlikely to work in high-premium firms.

Reduced-Form Evidence: Effect on Wages

DID estimation sheds light on the aggregate impact of the policy.

$$\begin{aligned}\ln y_{ijt} = \gamma \cdot Migrant_i \times Post_t + \delta \cdot Migrant_i + \lambda_t \\ + \beta_X \cdot X_{it} \times Post_t + \beta_{Id}\hat{\alpha}_i^d \quad (+\,\zeta_j) + \varepsilon_{it}\end{aligned}$$

- y_{ijt}: log wage of employee i in employer j in year t
- $Migrant_i$: $= 1$ for migrants
- $Post_t$: $= 1$ if in or after 2011, and $= 0$ otherwise
- X_{it}: observed employee characteristics, including age, age squared, and gender
- $\hat{\alpha}_i^d$: deciles of estimated person effect
- ζ_j: employer FE

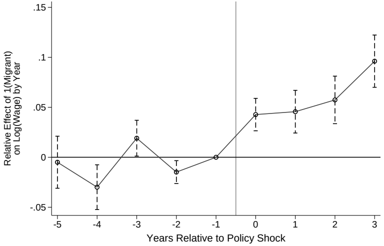
Reduced-Form Evidence: DID

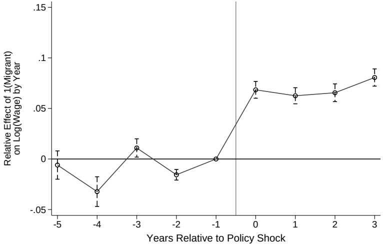
VARIABLES	(1) Ln(Income)	(2) Ln(Income)
Migrant	-0.123***	-0.0924***
	(0.000969)	(0.000482)
Migrant $\times$ Post	0.0636***	0.0761***
	(0.00107)	(0.000602)
Observations	5,553,049	5,553,049
R-squared	0.694	0.908
Year FE	Yes	Yes
Worker Controls	Yes	Yes
Worker Controls $\times$ Post	Yes	Yes
Employer FE	No	Yes
Employer FE $\times$ Post	No	Yes

- With employer FE: wage setting effect
- Without employer FE: wage setting effect + residual sorting effect

Reduced-form Evidence: Event Study

$$\ln y_{ijt} = \sum_{m=-4,m\neq -1}^{3} \gamma_m \cdot Migrant_i \times t_m + \delta \cdot Migrant_i + \lambda_t$$
$$+ \beta_X \cdot X_{it} \times Post_t + \beta^{Id} \hat{\alpha}_i^d \quad (+\zeta_j) + \varepsilon_{it}$$





Wage Setting Effect: Heterogeneity Analysis

$$\begin{aligned}\ln y_{ijt} =& \gamma_1 \cdot Migrant_i \times Treat_{ijt} \times Post_t + \gamma_2 \cdot Migrant_i \times Treat_{ijt} \\ &+ \gamma_3 \cdot Migrant_i \times Post_t + \gamma_4 \cdot Treat_{i,j} \times Post_t \\ &+ \beta_X \cdot X_{it} \times Post_t + \beta_{Id} \hat{\alpha}_i^d \quad (+\zeta_j) + \varepsilon_{it}\end{aligned}$$

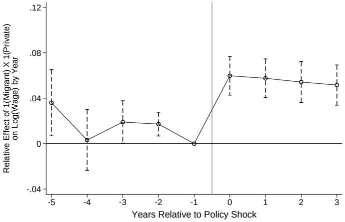
- $Treat_{ijt}$: different treatment groups:
  - Employees in the private sector
  - Employees with person effect (PE) above the median
  - Employees born in or after 1980

Wage Setting Effect: DDD

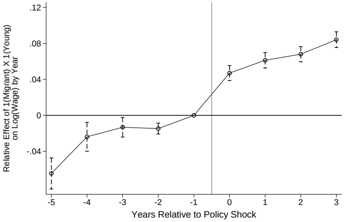
VARIABLES	(1) Log(Income)	(2) Log(Income)	(3) Log(Income)
Migrant	-0.124***	-0.0813***	-0.0391***
	(0.000912)	(0.000692)	(0.000638)
Migrant $\times$ Post	0.0475***	0.0589***	0.0432***
	(0.00114)	(0.000815)	(0.000852)
Migrant $\times$ Post $\times$ Private	0.0425***		
	(0.00133)		
Migrant $\times$ Post $\times$ High-skilled		0.0318***	
		(0.000944)	
Migrant $\times$ Post $\times$ Young Workers			0.0775***
			(0.000925)
Observations	5,553,049	5,553,049	5,553,049
R-squared	0.909	0.908	0.909
Worker Controls	Yes	Yes	Yes
Worker Controls $\times$ Post	Yes	Yes	Yes
Employer FE	Yes	Yes	Yes
Employer FE $\times$ Post	Yes	Yes	Yes

Wage Setting Effect: DDD Event Study

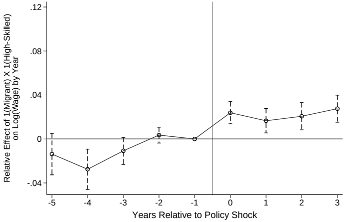
Figure: *Hukou* Quota Tightening and Wage Setting: Event Study



Private Sector



Young Workers



High-skilled Workers

Residual Sorting: Sectoral Choice of Workers

- Sectoral choice by workers' skills:

$$Private_{it} = \gamma_1 \cdot Migrant_i \times Post_t + \gamma_2 \cdot Migrant_i + \gamma_3 \cdot Skilled_i + \beta_X \cdot X_{it} \times Post_t + \beta_{Id}\hat{\alpha}_i^d + \varepsilon_{it}, \quad i \in \{Skilled_i = 0, 1\}$$

where $Skilled_i$: i's PE is above the 50th percentile

- Within high-skilled workers, explore the sectoral choice of workers by age:

$$Private_{it} = \gamma_1 \cdot Migrant_i \times Young_i \times Post_t + \gamma_2 \cdot Migrant_i \times Young_i + \gamma_3 \cdot Migrant_i \times Post_t + \gamma_4 \cdot Young_i^g \times Post_t + \beta_X \cdot X_{it} \times Post_t + \beta_{Id}\hat{\alpha}_i^d + \varepsilon_{it}, \quad i \in \{Skilled_i = 1\}$$

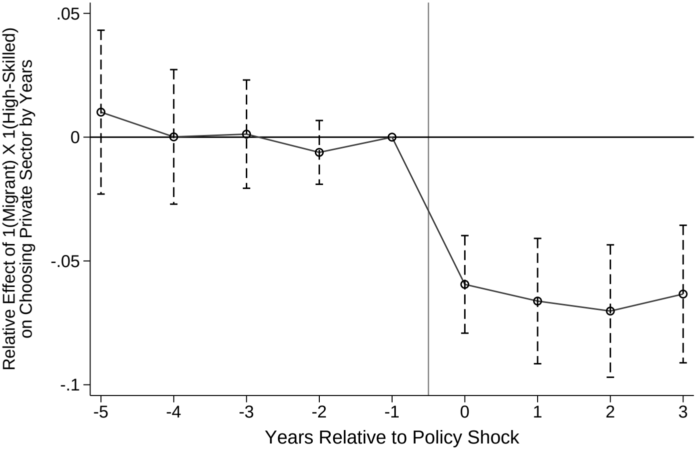
where $Young_i$: i is born after year 1980.

Residual Sorting: Sectoral Choice of Workers

- High-skilled migrants are less likely to work in the private sector after the policy shock, especially the younger ones.

	(1)	(2)	(3)
VARIABLES	Private	Private	Private
Sample	High-Skilled	Low-Skilled	High-Skilled
Migrant	0.160***	0.238***	0.131***
	(0.00190)	(0.00178)	(0.00254)
Migrant $\times$ Post	-0.0123***	0.0593***	-0.0140***
	(0.00170)	(0.00185)	(0.00214)
Migrant $\times$ Post $\times$ Young			-0.00647*
			(0.00339)
Observations	2,376,234	2,242,908	2,376,234
R-squared	0.117	0.200	0.119
Worker Controls	Yes	Yes	Yes
Worker Controls $\times$ Post	Yes	Yes	Yes

Residual Sorting: Sectoral Choice of Workers (High- vs Low-skilled)



Robustness

- Restrict employers and employees to present in the data for at least once during both the pre-shock and post-shock periods. Results
  - In baseline, we only restrict employers to present in the data for at least once during both the pre-shock and post-shock periods.
- Use alternative definitions of benchmark employers Results
  - In this robustness, we further require the within-employer wage gap between migrants and natives to be in the bottom quartile.
- Include job stayers Results
  - In baseline, we focus on job movers.
- Use only firms Results
  - In baseline, we include government agencies, schools, hospitals, etc.
- Evidence from a subsample with hukou information Results

Policy Implications: Value of a Hukou Quota

- The hukou quota's effect on wage setting is 0.051.
- Hukou quota before the reform is 0.2 (20% of migrants can receive hukou) and after the reform is 0.12 (40% reduction).
- For migrant workers staying in the same firm,

$$\begin{aligned} \ln w_{M1} - \ln w_{M0} &= -\ln(1 + a_{M1}c) + \ln(1 + a_{M0}c) \\ 0.037 &= -\ln(1 + 0.12c) + \ln(1 + 0.2c) \end{aligned}$$

- The value of hukou is $\ln(1 + c) = 0.40$, 40% of earnings.

Policy Implications: Potential Misallocation

- Employers in the public sector are less productive.
  1. Rich evidence in the literature about SOEs' low productivity
     - Berkowitz et al. (2017); Brandt et al. (2008); Brandt et al. (2022); Chen et al. (2021); Hsieh and Klenow (2009); Hsieh and Song (2015); Islam et al. (2006); Jefferson and Rawski (1994); Song et al. (2011)
  2. Investigation of the Annual Survey of Industrial Firms (ASIF) data
     - SOEs have lower profitability, MRPL, and TFP than POEs. Results
  3. Lower wage premium in terms of EE for native workers Results

Conclusions

- The migrant-native wage gap in the formal sector of a large city
  - Migrants have higher skills,
  - but lower wages given skill due to
    - wage setting
    - residual sorting
- Employer-sponsored *hukou* plays a role in the wage gap.
  - Employers with *hukou* quotas suppress wages for migrants by 40%.
  - Disproportionately affects young and skilled migrants.
- Policy implications
  - High-skilled workers sort into the public sector, which is granted more *hukou* quotas.
  - Potential misallocation: beneficial to assign the *hukou* quotas to high-productivity firms

Thank you for listening!

Appendix

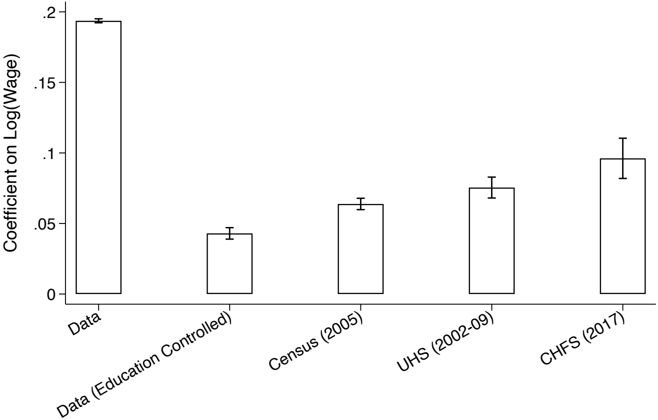
Observed Migration-Native Wage Gap ▸ Back

$$\ln y_{it} = \beta_0 + \beta_1 Migrant_i + \beta X_{it} + \gamma W_{jt} + \lambda_t + \varepsilon_{it}$$

- $\ln y_{git}$: the log-transformed wage of worker i in year t
- $Migrant_i = 1(0)$ if individual i is a migrant (native worker)
- β_1: the migrant-native wage gap
- X_{it}: observed characteristics of employees
- W_{jt}: observed characteristics of employer j
- λ_t: year fixed effect
- ε_{it}: the error term

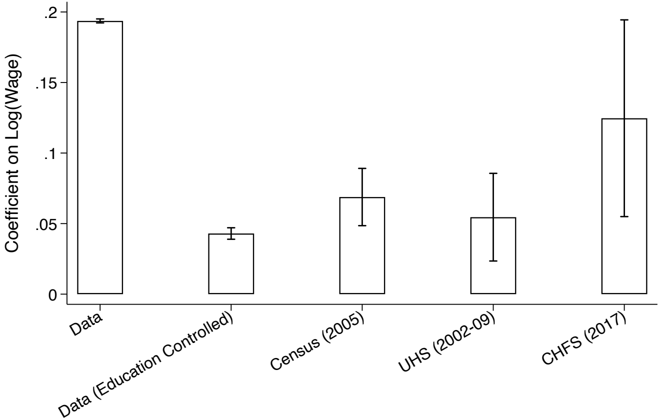
Observed Migrant-Native Wage Gap

Back



- For the three bars on the right, migrants are defined as individuals whose birth place is not in the surveyed city. We control for individual's gender, education, age, industry, occupation, etc.

Observed Migrant-Native Wage Gap ▸ Back



- This figure restricts the samples in the studied city. Other datasets show consistently positive migrant-native wage gap as in our data, despite the possible inclusion of informal sector in those datasets.

Data Representativeness ▸ Back

- There are more natives and more college-educated workers in the HPF data than in the aggregate data.
- Workers in the HPF data have higher income.

Table: Summary Statistics

	HPF Data	Aggregate Data	Source
Income (2006-2014)	69812.70	54674.72	Statistical Yearbook (07-15)
Income (2010)	66314.38	55462.36	Statistical Yearbook (11)
Female (2010)	0.46	0.43	Census 2010
Age (2010)	33.85	34.61	Census 2010
Native (2010)	0.59	0.45	Census 2010
College (2010)	0.47	0.27	Census 2010

Notes: For census data, only individuals aged between 22 to 50 years old and are currently working are included.

Heterogeneous Wage Gaps

Back

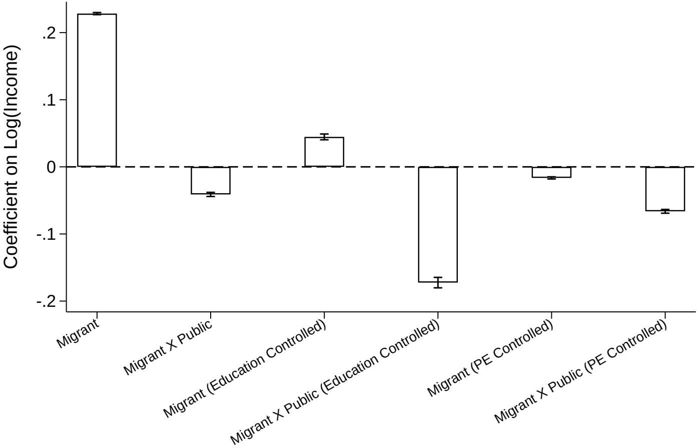
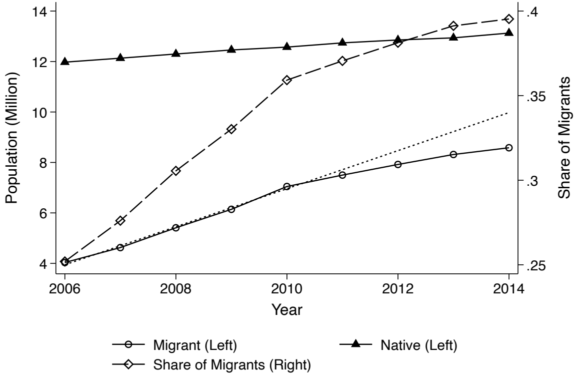


Figure: Migrant-Native Wage Gap

Labor Supply in the City ▸ Back

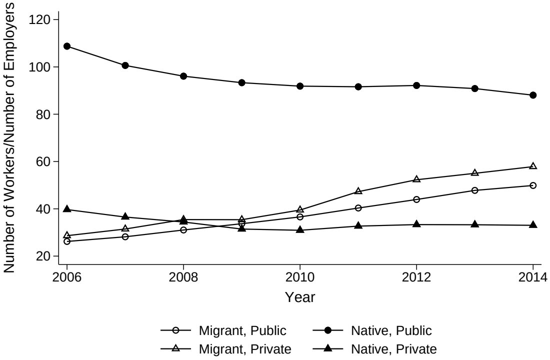
- While the growth of migrant inflow slowed, it still increased over the sample period.



Source: Municipal Statistical Yearbooks. Here migrants are defined as non-local residents.

Labor Supply in the City ▸ Back

- The average number of migrants in each firm is increasing as well.
- The increase is even larger for the private sector after 2011. If anything, the labor supply story should work against the *hukou* channel.



Source: HPF data.

Migrants with *Hukou*: Census 2010 Data

Back

- Here, high-skilled is defined as having any college education.

	Without Hukou	With Hukou
Migrant, Total	79.7%	20.4%
Migrant, Low-Skilled	92.9%	7.1%
Migrant, High-Skilled	58.7%	41.3%
Native, Total	0.6%	99.4%
Native, Low-Skilled	0.7%	99.3%
Native, High-Skilled	0.6%	99.4%

Appendix

App.I. Descriptive Analysis

App.II. Results on the AKM model

App.III. Misallocation of Workers

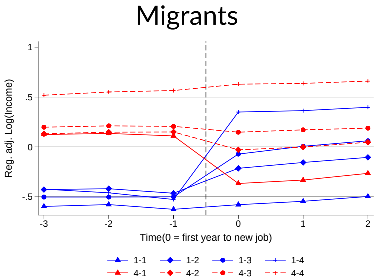
App.IV. Robustness

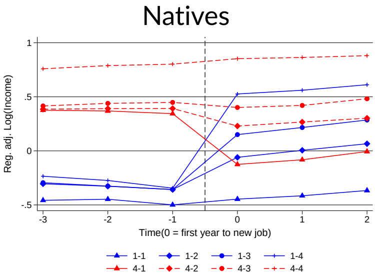
Assumptions ▸ Back

- Model specification assumption
  - Exogenous mobility assumption: no sorting based on unobserved comparative advantage
  - Separability assumption: additive separable PE and EE
  - Fixed effects are identified by job movers in each connected set.
- Testable implications
  - Symmetric wage changes for different job switches
  - No pre-trend before job switches
  - No systematic patterns of unexplained residuals
  - High explanatory power of the model to wage variations

Validation of AKM: Event Studies

Back



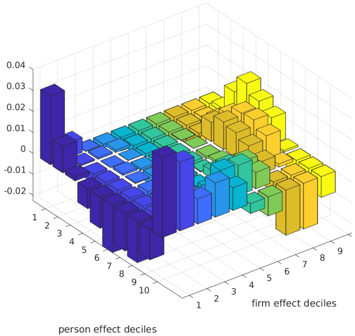


- No pre-trends in wage
- No effect for moves within the same EE quartile
- Symmetric wage changes in opposite moves

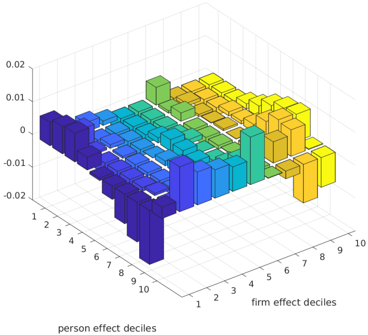
Validation of AKM: Residuals

Back

Migrants



Natives



- Overall small residuals; uncorrelated with PE & EE deciles

Decomposition of Wage Variation ▸ Back

$$\begin{aligned} Var(y_{it}) =& Var(\alpha_i) + Var(\Psi_J) + Var(X'_{it} \cdot \beta) \\ &+ 2Cov(\alpha_i, \Psi_J) \\ &+ 2Cov(\alpha_i, X'_{it} \cdot \beta) + 2Cov(\Psi_J, X'_{it} \cdot \beta) \\ &+ Var(r_{it}) \end{aligned}$$

Decomposition of Wage Variation ▸ Back

	Sample Chinese city '06-'14 Migrants	Sample Chinese city '06-'14 Natives	W. Germany '02-'09	U.S '07-'13	Portugal '02-'09 Males	Brazil '02-'14 White males
SD of log wages	0.797	0.745	0.499	0.961	0.554	0.670
Mean log wages	11.078	10.843				
AKM decomposition						
SD of person effect	0.475	0.500	0.357	0.690	0.420	0.484
SD of employer effect	0.427	0.348	0.230	0.285	0.247	0.304
SD of $X'_{it} \cdot \beta$	0.346	0.292	0.084	0.059	0.069	0.175
Corr. btw PE & EE	0.167	0.155	0.249	0.232	0.167	0.275
Adj. R^2	0.873	0.881	0.927		0.934	0.901
Adj. R^2 with match effect	0.911	0.912				
% of log wage variance due to						
Person effect	35.5	45.2	51.2	51.5	57.6	52.1
Employer effect	28.7	21.9	21.2	8.7	19.9	20.6
Cov. btw PE & EE	9.6	9.6	24.9	11.7	11.4	18.0
EE + cov(PE,EE)	38.3	31.4	46.1	20.4	31.3	38.6
in largest connected set						
# of employers (mil)	0.035	0.037			0.21	0.18
# of movers (mil)	0.67	0.65			1.89	3.55
# of person-year obs. (mil)	2.9	3.5			8.2	22
Source			CHK'13 Tbl. III	SPGBW'19 Tbl. III	CCK'16 Tbl. II	GLSC'21 Tbl. 2

AKM Results ▸ Back

	Native, pre-shock (1)	Migrant, pre-shock (2)	Native, post-shock (3)	Migrant, post-shock (4)
Largest connected set				
Standard deviation of log wages	0.729	0.778	0.717	0.777
Mean log wages	10.680	10.902	11.025	11.242
Variance decomposition				
SD of person effects	0.531	0.494	0.534	0.496
SD of employer effects	0.369	0.458	0.343	0.432
SD of covariates	0.240	0.289	0.203	0.241
Correlation of person/employer effects	0.066	0.070	0.089	0.121
Adjusted R^2 of model	0.894	0.885	0.928	0.919
Adjusted R^2 with match effect	0.909	0.904	0.942	0.939
Percentage of variance of log wages due to:				
Person effect	53.0	40.3	55.4	40.8
Employer effect	25.6	34.6	22.9	30.9
Covariance of person and employer effects	7.2	7.1	4.7	5.0
Emp. effects + covariance person and emp. effects	32.8	41.7	27.5	35.9
Number of employers	29998	27406	28684	27480
Number of movers	533219	468471	549967	578449
Number of person-year observations	1768352	1299231	1653029	1561127

$$\mathrm{corr}(PE^M_{pre}, PE^M_{post}) = 0.697, \quad \mathrm{corr}(PE^N_{pre}, PE^N_{post}) = 0.773$$

$$\mathrm{corr}(EE^M_{pre}, EE^M_{post}) = 0.619, \quad \mathrm{corr}(EE^N_{pre}, EE^N_{post}) = 0.629$$

Limited Mobility Bias

- AKM models identified by the “connected set”
  - Employers connected by job switches
  - Sample usually includes the largest connected set
- Some employers are “thinly” connected due to “limited mobility”
  - Variance of EE *upward* biased
  - Correlation between PE and EE *downward* biased
- Leave-one-out estimation
  - Kline, Saggio, Sølvsten (KSS, 2020)
  - Connected sample becomes smaller

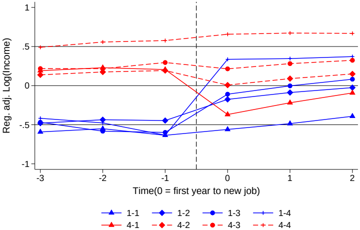
KSS Variance Decomposition

	Native, pre-shock (1)	Migrant, pre-shock (2)	Native, post-shock (3)	Migrant, post-shock (4)
Leave-one-out connected set				
Standard deviation of log wages	0.693	0.721	0.660	0.704
Mean log wages	10.706	11.147	10.966	11.314
Variance decomposition				
SD of person effects	0.423	0.376	0.331	0.382
SD of employer effects	0.329	0.425	0.316	0.413
Correlation of person/employer effects	0.195	0.214	0.251	0.214
Adjusted R^2 of model	0.709	0.752	0.601	0.775
Percentage of variance of log wages due to:				
Person effect	37.2	27.3	25.2	29.4
Employer effect	22.5	34.7	22.9	34.4
Covariance of person and employer effects	11.3	13.2	12.0	13.6
Emp. effects + covariance person and emp. effects	33.8	47.9	34.9	48.0
Number of employers	22344	19611	21431	20442
Number of movers	217462	168875	162813	205333
Number of person-year observations	888448	615659	560747	670267

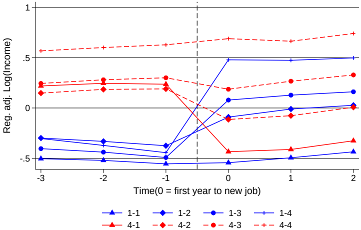
Validation of AKM on Four Groups

Back

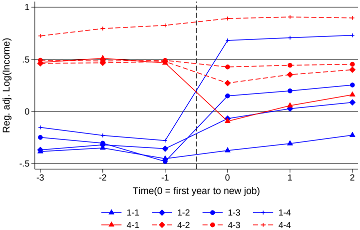
Migrants, Pre-shock



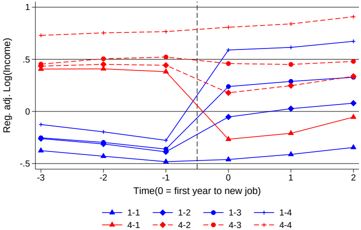
Migrant, Post-shock



Native, Pre-shock



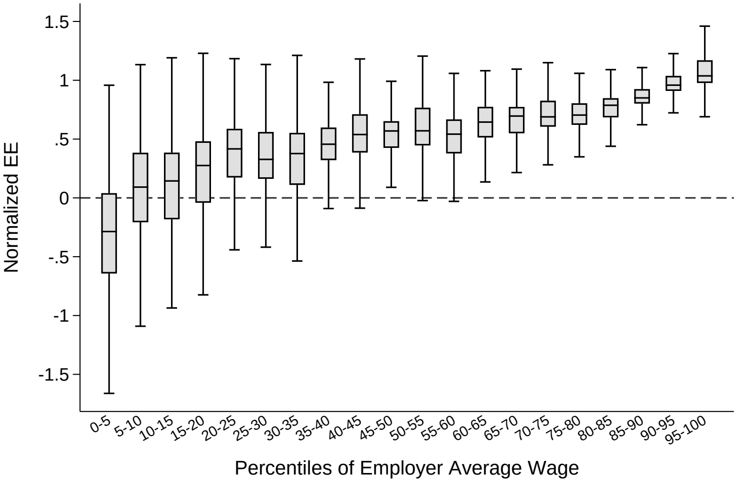
Native, Post-shock



Procedure of Normalization ▸ Back

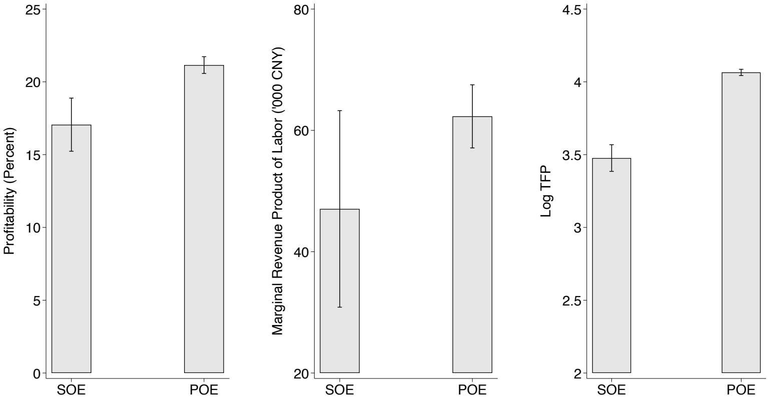
1. Obtain four sets of estimated PE and EE with AKM model estimations (before vs. after, migrants vs. natives).
2. Define the benchmark employers:
   - Categorize employers into percentiles based on the average wage before and after the policy shock.
   - Identify private-sector employers whose average wage falls within the bottom 5-10th percentile in each period.
   - Define benchmark employers as the intersection of these two groups of employers.
3. Calculate the mean of the estimated EE for benchmark employers in each period (before vs. after) separately for migrants and natives.
4. Subtract the estimated EE of other employers from the corresponding mean of the benchmark EE.
5. Add the estimated PE for workers using the corresponding mean of the benchmark EE.

Normalized EE and Average Wage ▸ Back



Appendix

Profitability, MRPL, and TFP of Firms in ASIF



Profitability, MRPL, and TFP of Firms in ASIF ▸ Back

- Profitability: Total profits over net value of fixed assets, following Song et al. (2011).
- MRPL:
  - Assuming a Cobb-Douglas production function

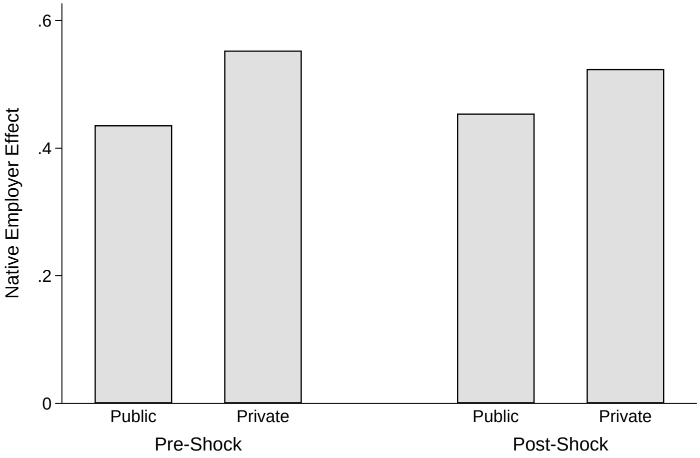
$$Y_j = A_j L_j^{\beta_L} K_j^{\beta_K}$$

  - Estimate the production function with the Ackerberg-Caves-Frazer method (Ackerberg et al., 2015)
  - Then,

$$MRPL_j = \beta_L \frac{pY_j}{L_j}$$

- TFP: with production function estimation when calculating MRPL, we have the estimation of $\log(A_j)$.

Employer Effect for Native Workers ▸ Back



Appendix

Restricted Sample ▸ Back

	(1) Pre-shock	(2) Post-shock	(3) Difference (Post-Pre)
Wage Gap (Migrant - Native, same below)	0.197	0.276	0.079
1. Person Effect Gap	0.329	0.340	0.010
2. Covariates Gap	-0.112	-0.052	0.060
3. Employer Effect Gap	-0.021	-0.012	0.009
3.1. Wage Setting Effect	-0.033	-0.009	0.024
3.2. Sorting Effect	0.013	-0.003	-0.016
3.2.1. Skill-Based Sorting Effect	0.062	0.054	-0.008
3.2.2. Residual Sorting Effect	-0.049	-0.057	-0.008

Restricted Sample

Back

VARIABLES	(1) Log(Income)	(2) Log(Income)	(3) Log(Income)	(4) Log(Income)
Migrant	-0.127***	-0.0551***	-0.0395***	-0.0895***
	(0.000846)	(0.000717)	(0.000667)	(0.000452)
Migrant × Post	0.0416***	0.0521***	0.0341***	0.0675***
	(0.00115)	(0.00101)	(0.000929)	(0.000606)
Migrant × Post × Private	0.0396***			
	(0.00128)			
Migrant × Post × High-skilled		0.0275***		
		(0.00120)		
Migrant × Post × Young Workers			0.0736***	
			(0.00111)	
Migrant × Post × Newly-hired				0.0308***
				(0.00317)
Observations	4,802,578	4,802,578	4,802,578	4,802,578
R-squared	0.900	0.899	0.900	0.899
Worker Controls	Yes	Yes	Yes	Yes
Worker Controls × Post	Yes	Yes	Yes	Yes
Employer Controls	Yes	Yes	Yes	Yes
Employer Controls × Post	Yes	Yes	Yes	Yes

Restricted Sample

Back

VARIABLES	(1) Wage Setting	(2) Wage Setting	(3) Residual Sorting	(4) Residual Sorting
Post	0.00752*	-0.0151**	0.0559	-0.0785
	(0.00409)	(0.00591)	(0.0779)	(0.0516)
Private Sector $\times$ Post		0.0362***		0.244**
		(0.00912)		(0.115)
Observations	119,750	119,750	109,741	109,741
R-squared	0.014	0.061	0.011	0.025
Employer Controls	Yes	Yes	Yes	Yes

Restricted Sample ▸ Back

VARIABLES	(1) Private	(2) Private	(3) Private
Migrant	0.166***	0.245***	0.134***
	(0.00192)	(0.00175)	(0.00257)
Migrant $\times$ Post	-0.0129***	0.0546***	-0.0134***
	(0.00173)	(0.00181)	(0.00220)
Migrant $\times$ Post $\times$ Young			-0.00832**
			(0.00347)
Observations	2,276,711	2,259,170	2,276,711
R-squared	0.102	0.196	0.103
Sample	High-Skilled	Low-Skilled	High-Skilled
Worker Controls	Yes	Yes	Yes

Alternative Benchmark Employers ▸ Back

Baseline definition + within-employer wage gap in the bottom quartile

	(1) Pre-shock	(2) Post-shock	(3) Difference (Post-Pre)
Wage Gap (Migrant - Native, same below)	0.194	0.205	0.011
1. Person Effect Gap	0.326	0.277	-0.049
2. Covariates Gap	-0.120	-0.069	0.051
3. Employer Effect Gap	-0.012	-0.003	0.009
3.1. Wage Setting Effect	-0.026	0.011	0.037
3.2. Sorting Effect	0.014	-0.014	-0.028
3.2.1. Skill-Based Sorting Effect	0.064	0.051	-0.014
3.2.2. Residual Sorting Effect	-0.051	-0.065	-0.014

Alternative Benchmark Employers

Back

Baseline definition + within-employer wage gap in the bottom quartile

VARIABLES	(1) Log(Income)	(2) Log(Income)	(3) Log(Income)	(4) Log(Income)
Migrant	-0.143***	-0.0735***	-0.0578***	-0.109***
	(0.000811)	(0.000686)	(0.000626)	(0.000429)
Migrant × Post	0.0442***	0.0483***	0.0374***	0.0720***
	(0.00103)	(0.000836)	(0.000820)	(0.000518)
Migrant × Post × Private	0.0399***			
	(0.00114)			
Migrant × Post × High-skilled		0.0415***		
		(0.000998)		
Migrant × Post × Young Workers			0.0798***	
			(0.000965)	
Migrant × Post × Newly-hired				0.0217***
				(0.00256)
Observations	5,546,655	5,546.655	5,546,655	5,546,655
R-squared	0.904	0.903	0.904	0.903
Worker Controls	Yes	Yes	Yes	Yes
Worker Controls × Post	Yes	Yes	Yes	Yes
Employer Controls	Yes	Yes	Yes	Yes
Employer Controls × Post	Yes	Yes	Yes	Yes

Alternative Benchmark Employers

Back

Baseline definition + within-employer wage gap in the bottom quartile

VARIABLES	(1) Wage Setting	(2) Wage Setting	(3) Residual Sorting	(4) Residual Sorting
Post	0.00433	-0.0222***	0.0480	-0.0723
	(0.00403)	(0.00685)	(0.0738)	(0.0533)
Private Sector × Post		0.0420***		0.207**
		(0.0100)		(0.105)
Observations	129,441	129,441	118,643	118643
R-squared	0.024	0.075	0.010	0.023
Employer Controls	Yes	Yes	Yes	Yes

Alternative Benchmark Employers ▸ Back

Baseline definition + within-employer wage gap in the bottom quartile

VARIABLES	(1) Private	(2) Private	(3) Private
Migrant	0.162***	0.245***	0.129***
	(0.00192)	(0.00177)	(0.00257)
Migrant $\times$ Post	-0.00814***	0.0590***	-0.00738***
	(0.00171)	(0.00184)	(0.00216)
Migrant $\times$ Post $\times$ Young			-0.0123***
			(0.00341)
Observations	2,373,370	2,245,772	2,373,370
R-squared	0.103	0.195	0.104
Sample	High-Skilled	Low-Skilled	High-Skilled
Worker Controls	Yes	Yes	Yes

Include Job Stayers ▸ Back

	(1) Pre-shock	(2) Post-shock	(3) Difference (Post-Pre)
Wage Gap (Migrant - Native, same below)	0.134	0.047	-0.086
1. Person Effect Gap	0.278	0.158	-0.120
2. Covariates Gap	-0.097	-0.061	0.036
3. Employer Effect Gap	-0.048	-0.050	-0.002
3.1. Wage Setting Effect	-0.051	-0.002	0.049
3.2. Sorting Effect	0.003	-0.048	-0.051
3.2.1. Skill-Based Sorting Effect	0.048	0.020	-0.028
3.2.2. Residual Sorting Effect	-0.045	-0.068	-0.023

Include Job Stayers ▸ Back

VARIABLES	(1) Log(Income)	(2) Log(Income)	(3) Log(Income)	(4) Log(Income)
Migrant	-0.0906***	-0.0112***	-0.0274***	-0.0583***
	(0.000453)	(0.000459)	(0.000413)	(0.000290)
Migrant × Post	0.0399***	0.0527***	0.0239***	0.0707***
	(0.000549)	(0.000533)	(0.000528)	(0.000340)
Migrant × Post × Private	0.0523***			
	(0.000664)			
Migrant × Post × High-skilled		0.0254***		
		(0.000663)		
Migrant × Post × Young Workers			0.102***	
			(0.000660)	
Migrant × Post × Newly-hired				0.0378***
				(0.00183)
Observations	17,591,404	17,591,404	17,591,404	17,591,404
R-squared	0.904	0.903	0.903	0.903
Worker Controls	Yes	Yes	Yes	Yes
Worker Controls × Post	Yes	Yes	Yes	Yes
Employer Controls	Yes	Yes	Yes	Yes
Employer Controls × Post	Yes	Yes	Yes	Yes

Include Job Stayers ▸ Back

VARIABLES	(1) Wage Setting	(2) Wage Setting	(3) Residual Sorting	(4) Residual Sorting
Post	0.0427***	0.0207***	-0.0191**	-0.0526***
	(0.00452)	(0.00624)	(0.00800)	(0.0106)
Private Sector × Post		0.0388***		0.0353**
		(0.00861)		(0.0149)
Observations	132,248	132,248	121,209	121,209
R-squared	0.049	0.063	0.005	0.074
Employer Controls	Yes	Yes	Yes	Yes

Include Job Stayers ▸ Back

VARIABLES	(1) Private	(2) Private	(3) Private
Migrant	0.181***	0.218***	0.140***
	(0.00111)	(0.00119)	(0.00141)
Migrant × Post	0.00584***	0.0485***	0.000604
	(0.000871)	(0.00113)	(0.00101)
Migrant × Post × Young			-0.0113***
			(0.00178)
Observations	7,046,087	6,273,658	7,046,087
R-squared	0.120	0.196	0.123
Sample	High-Skilled	Low-Skilled	High-Skilled
Worker Controls	Yes	Yes	Yes

With Only Firms ▸ Back

VARIABLES	(1) Log(Income)	(2) Log(Income)	(3) Log(Income)	(4) Log(Income)
Migrant	0.0582***	0.103***	0.115***	0.0777***
	(0.00117)	(0.000818)	(0.000719)	(0.000522)
Migrant × Post	0.0251***	0.0283***	0.0162***	0.0320***
	(0.00172)	(0.00116)	(0.00103)	(0.000714)
Migrant × Post × Private	0.0120***			
	(0.00182)			
Migrant × Post × High-skilled		0.00290**		
		(0.00140)		
Migrant × Post × Young Workers			0.0520***	
			(0.00129)	
Migrant × Post × Newly-hired				0.00405
				(0.00346)
Observations	3,963,702	4,314,818	4,314,818	4,314,818
R-squared	0.903	0.898	0.898	0.898
Worker Controls	Yes	Yes	Yes	Yes
Worker Controls × Post	Yes	Yes	Yes	Yes
Employer Controls	Yes	Yes	Yes	Yes
Employer Controls × Post	Yes	Yes	Yes	Yes

With Only Firms ▸ Back

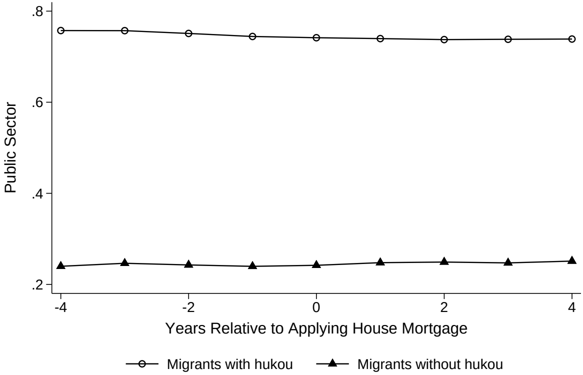
VARIABLES	(1) Wage Setting	(2) Wage Setting	(3) Residual Sorting	(4) Residual Sorting
Post	0.0293***	-0.0156	0.0893	-0.356***
	(0.00521)	(0.0109)	(0.114)	(0.119)
Private Sector × Post		0.0518***		0.583**
		(0.0143)		(0.292)
Observations	104,889	104,889	93,728	93,728
R-squared	0.058	0.101	0.010	0.041
Employer Controls	Yes	Yes	Yes	Yes

With Only Firms ▸ Back

VARIABLES	(1) Private	(2) Private	(3) Private
Migrant	0.154***	0.222***	0.158***
	(0.00191)	(0.00177)	(0.00240)
Migrant $\times$ Post	-0.0504***	0.0267***	-0.0501***
	(0.00169)	(0.00174)	(0.00207)
Migrant $\times$ Post $\times$ Young			0.00531
			(0.00354)
Observations	2,036,937	1,991,986	2,036,937
R-squared	0.088	0.201	0.089
Sample	High-Skilled	Low-Skilled	High-Skilled
Worker Controls	Yes	Yes	Yes

Evidence with *Hukou* Information

- For a 7% sample that applied for house mortgage, we observe their *hukou* status at the time of mortgage application.
- Migrants who obtained a *hukou* $\rightarrow$ 80% in the public sector
- Migrants who did not obtain a *hukou* $\rightarrow$ 80% in the private sector



Evidence with *Hukou* Information

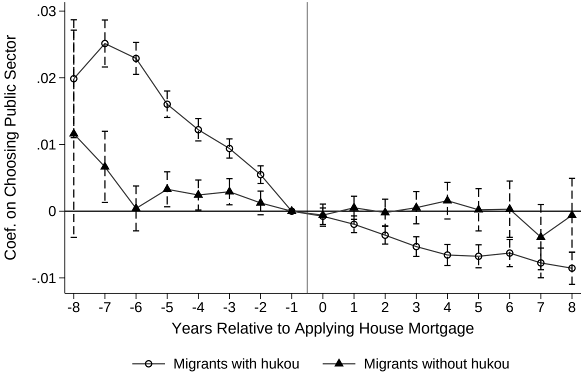
- The dynamics of sector choice of those who obtained a local *hukou* and did not obtain one should differ as well.
- Two separate "event studies" on sector choice: comparing migrants with vs. without a *hukou*

$$Public_{it} = \alpha_i + \sum_{m=-8, m\neq -1}^{8} \gamma_m t_m + X'_{it}\beta_X + \varepsilon_{it}, \quad i \in \{Hukou_i = 0, 1\}$$

  - $Hukou_i$: =1 when a migrant i has obtained a local hukou, = 0 when a migrant i has not
  - t_m: year relative to the mortgage application

Evidence with *Hukou* Information

- Migrants who did not obtain a hukou: coefficients are insignificant
- Migrants who obtained a hukou: leaving the public sector



Evidence with *Hukou* Information

- We compare natives with: 1) migrants that have already obtained *hukou* before the policy shock, and 2) migrants that haven't even afterward.
- Basic idea: migrants who have obtained *hukou* should be largely "immune" to the policy shock.

$$\begin{aligned}\ln y_{it} = & \alpha_i + \gamma_1 \cdot g_i \times Hukou\_pre_i \times Post_t + \\ & \gamma_2 \cdot g_i \times No\_Hukou\_post_i \times Post_t + \\ & \gamma_3 \cdot g_i \times Hukou\_pre_i + \gamma_4 \cdot Hukou\_pre_i \times Post_t + \\ & \gamma_5 \cdot g_i \times No\_Hukou\_post_i + \gamma_6 \cdot No\_Hukou\_post_i \times Post_t + \\ & X_{it}' \beta_X + \varepsilon_{it}\end{aligned}$$

Evidence with *Hukou* Information ▸ Back

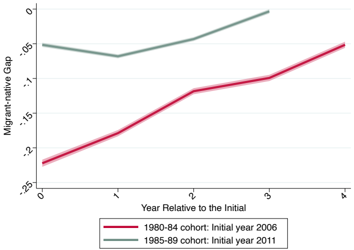
VARIABLES	(1) Ln(Income)	(2) Ln(Income)
Migrant × Post × With Hukou Before Policy	0.0245***	0.0428***
	(0.00414)	(0.00225)
Migrant × Post × No Hukou After Policy	0.132***	0.0824***
	(0.0198)	(0.00820)
Observations	701,088	698,960
R-squared	0.689	0.907
Sample	Mortgage Applier	Mortgage Applier
Individual Controls	Yes	Yes
Individual Controls × Post	Yes	Yes
Employer FE	No	Yes
Employer FE × Post	No	Yes

Dynamics of Wage Gap

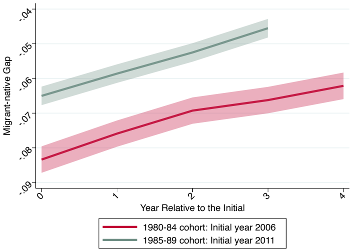
$$\text{ln} w_{igt} = \beta_t Mig_i + \rho_d \hat{\alpha}_i^d + \lambda_{ct} + \varepsilon_{igt}$$

- Run the regression separately for each year t.
- $\hat{\alpha}_i^d$ is the deciles of person effect.
- λ_{ct} control for gender-birth year.
- Replace log wage with employer effect to look at how workers move up the job ladder.
- Plot β_t.

Dynamics of Migrant-native Wage Gap

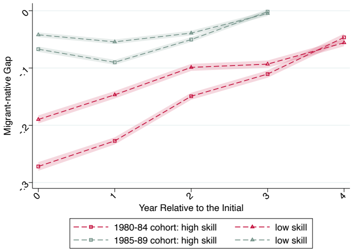


(a) Log Wage Gap

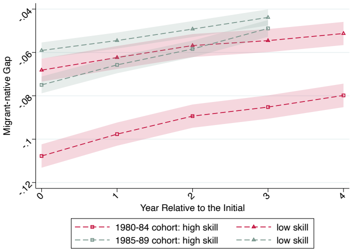


(b) Employer Effect Gap

Dynamics of Migrant-native Wage Gap: by Skill

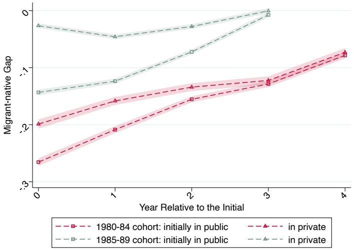


(a) Log Wage Gap

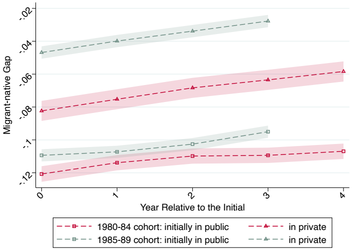


(b) Employer Effect Gap

Dynamics of Migrant-native Wage Gap: by Sector in the Initial Period



(a) Log Wage Gap



(b) Employer Effect Gap

Employer-Level Evidence

- Employer-level regression on the decomposed components of the migrant-native wage gap
- Two dependent variables (scaled by the average of firm share):

$$Wage\_Setting_{jT(t)} = (\psi^{M}_{jT(t)} - \psi^{N}_{jT(t)}) \cdot \pi^{M}_{jT(t)} / \overline{\pi_{jT(t)}}$$

$$Resid\_Sorting_{jT(t)} = \psi^{N}_{jT(t)} \left[\left(\pi^{M}_{jT(t)} - \pi^{M*}_{jT(t)} \right) - \left(\pi^{N}_{jT(t)} - \pi^{N*}_{jT(t)} \right) \right] / \overline{\pi_{jT(t)}}$$

VARIABLES	(1) Wage Setting	(2) Residual Sorting
Post	0.00583	-0.0812***
	(0.00957)	(0.0189)
Private Sector $\times$ Post	0.0208**	0.104***
	(0.00989)	(0.0307)
Observations	129,441	118,643
R-squared	0.019	0.006
Employer Controls	Yes	Yes