"Migrants and Firms: Evidence from China"

Imbert et al. (AER 2022)

Ben Pirie

ECON 232M

Spring 2022

Introduction

Research Question: How does rural-urban migration shape urban production in developing countries?

Method:

- Use longitudinal data on Chinese manufacturing firms to measure outcomes
- Rely on a shift-share design for identification

Motivation:

- Firm responses to factor supply shocks
- Effect of immigration on labor markets
- Important insight for developing economies

More Motivation

It's well understood that firms in developing countries have lower productivity per worker. This has been explained by a variety of factors:

- Imperfect access to capital (Banerjee and Duflo, 2014), inputs (Boehm and Oberfield, 2020), or technology (Howitt, 2000)
- Isolation from international markets (Verhoogen, 2008)
- Poor management practices Bloom et al., 2013; Atkin et al., 2017)
- **Abundance of cheap migrant labor (this paper)**

Firm Data

Source: Annual Chinese manufacturing census from the National Bureau of Statistics (2000 – 2006)

- all state-owned and non-state "firms" with sales exceeding RMB 5 million ($600,000) distribution
- accounts for 90% of total manufacturing output

Compose a balanced panel of 31,886 firms which exist in each annual census

- Includes basic firm data: location, industry, ownership type, exporting activity, number of employees
- Also includes a range of accounting information: sales, inputs, value added, wage bill, etc.

Firm Data

TABLE 1—SUMMARY STATISTICS OF KEY FIRM-LEVEL OUTCOMES.

	Balanced sample of firms		Unbalanced sample of firms	
	2000	2006–2000	2000	2006–2000
Labor cost	2.282	0.572	2.170	0.615
Employment	4.901	0.261	4.406	0.040
K/L ratio	3.802	0.254	3.701	0.069
Y/L ratio	3.680	0.565	3.570	0.638
Immigration rate	-	0.329	-	0.316
Observations	31,886	31,886	79,980	454,781

Sources: NBS firm-level data (2000, 2006). This table presents the baseline value in 2000 and growth between 2000–2006 for the key outcome variables. The first and second columns report statistics based on the balanced sample of firms; the third and fourth columns report statistics based on all firms present in the NBS data. *Labor cost* is the (log) compensation per worker including social security and housing benefits. *Employment* is the (log) number of workers. *K/L ratio* is the (log) ratio of fixed assets to employment. *Y/L ratio* is the (log) ratio of value added to employment. *Immigration rate* is the ratio of immigrants between 2000 and 2005 to the number of urban residents at destination in 2000.

Migration Data

Source: representative 1% Population Survey from 2005 from the National Bureau of Statistics

- observe household registration type (agricultural or non-agricultural), place of registration, and place of residence
- migrants are asked the main reason for leaving, the year they left, and their place of residence one and five years prior

Create a matrix of annual rural-urban migration flows between all Chinese prefectures from 2000 to 2005

- 45 million rural workers migrated to cities in this period (15% of the urban population in 2000)
- include migrants 15-64 years old
- exclude students or those migrating for school (¡ 5% of total)

Empirical Challenge

Goal: identify the adjustment of production to migrant inflows in cities

Problem: relocation of workers is endogenous

- migration is a function of labor demand and labor supply shocks

Solution: Bartik, or "shift-share" design

- combine origin labor supply shocks ("shifts") with historical migration patterns ("shares") into an instrument for migration flows, z_d

$$z_d = \sum_{o \in \Theta \setminus \{d\}} \lambda_{od} s_o$$

"Shift" Component

1. Compute the time-invariant potential agricultural output for a given crop, c, and prefecture, o

$$q_{co} = h_{co} \times y_{co}$$

where h_{co} is the total harvested area and y_{co} is the potential yield per hectare map

2. Construct international crop prices as the average price across all countries (excluding China), weighted by their baseline share in global exports

$$\log(p_{ct}) = \theta \log(p_{ct-1}) + \eta_t + \nu_c + \varepsilon_{ct}$$

where the residual, $\hat{\varepsilon}_{ct}$, is interpreted as year $\times$ crop innovation in log-prices.

"Shift" Component (Cont.)

Combine crop prices with crop patterns to construct the "shift"

$$s_{ot} = \left(\sum_c \bar{p}_c q_{co} \hat{\varepsilon}_{ct} \right) / \left(\sum_c \bar{p}_c q_{co} \right)$$

where $\bar{p}_c$ is the nominal international price for crop c at the baseline and $\hat{\varepsilon}_{ct}$ captures variation in world demand and supply over time.

"Share" Component

1. Compute the emigration rate for a given prefecture, o, and year, t

$$n_{ot} = \frac{\sum_{d \in \Theta \setminus o} M_{odt}}{P_o}$$

2. Compute the immigration rate for a given prefecture, d, and year, t

$$m_{dt} = \frac{\sum_{o \in \Theta \setminus d} M_{odt}}{P_d}$$

3. Historical settlement patterns between origin, o, and destination, d, are captured as the stock of migrants in 2000

$$\lambda_{od} = \frac{\sum_{t<2000} M_{odt}}{\sum_d \sum_{t<2000} M_{odt}}$$

Key Details

The validity of the shift-share instrument relies on two facts:

- agricultural returns affect emigration from rural areas

$$n_o = \beta_0 + \beta_1 s_o + \eta_o$$

- the instrument z_o predicts immigration flows

$$m_d = \alpha_0 + \alpha_1 z_d + \eta_d$$

First-Stage Results

TABLE 2—ORIGIN-BASED MIGRATION PREDICTIONS.

	Emigration	
	Inter-prefecture	Outside 300-km radius
Panel A: Emigration rate from origin		
Price shock	-0.117	-0.045
	(0.019)	(0.019)
Observations	335	335
	Immigration	
	Inter-prefecture	Outside 300-km radius
Panel B: Immigration rate at destination		
Shift-share instrument	-1.620	-1.305
	(0.425)	(0.368)
Observations	315	315

Notes: Robust standard errors are reported between parentheses. In Panel A, the dependent variable is the number of rural emigrants to urban areas in other prefectures or in prefectures located outside of a 300-km radius around the origin, divided by the number of rural residents at origin. In Panel B, the dependent variable is the number of rural immigrants from other prefectures or prefectures located outside of a 300-km radius around the destination divided by the number of urban residents at destination. See Section I and Equations (2) and (4) for a more comprehensive description of the two specifications.

Empirical Strategy

Baseline specification:

$$\Delta y_{id} = \alpha + \beta m_d + \varepsilon_{id}$$

where Δy_{id} is the change in outcome between 2000 and 2006 for firm i in prefecture d.

- Use 2SLS with the shift-share variable, z_d, as an instrument for immigration, m_d.

Identification relies on three conditions:

(i) shifts are quasi-randomly assigned across origins

(ii) there are many uncorrelated shifts

(iii) exclusion restriction holds

Employment Response to Migration

TABLE 3—IMPACT OF MIGRATION INFLOWS ON URBAN FIRMS.

	Labor cost (1)	Employment (2)	K/L ratio (3)	Y/L ratio (4)
Panel A: OLS estimates				
Migration	-0.172	0.270	-0.269	-0.339
	(0.057)	(0.035)	(0.044)	(0.071)
Observations	31,886	31,886	31,886	31,886
Panel B: IV estimates				
Migration	-0.147	0.294	-0.431	-0.437
	(0.062)	(0.053)	(0.095)	(0.108)
Observations	31,886	31,886	31,886	31,886
F-stat. (first stage)	23.59	23.59	23.59	23.59

Notes: Standard errors are clustered at the prefecture level and reported between parentheses. The sample is composed of the firms present every year in the NBS firm census between 2000 and 2006. *Migration* is the immigration rate, i.e., the migration flow divided by population at destination and at baseline. *Labor cost* is the (log) compensation per worker including social security and housing benefits. *Employment* is the (log) number of workers. *K/L ratio* is the (log) ratio of fixed assets to employment. *Y/L ratio* is the (log) ratio of value added to employment. See Section I and Equation (5) for a description of the IV specification.

Key Takeaway: firms choose not to mechanize due to the availability of cheap labor.

Employment Response to Migration (Cont.)

- Wages decline, though not as much as predicted by other literature on internal migration in developing economies.
- A large portion of migrant workers are not hired by the firms in their sample.
- The decline in K/L suggests capital does not adjust with increased employment.
  - (i) high-substitutability between labor and capital?
  - (ii) credit constraints or adjustment costs?
  - (iii) change in production line organization and adoption of new technologies with different factor intensities?
- The decline in Y/L suggests that workers are becoming less productive.

robustness

Production Response to Migration

TABLE 7—IMPACT OF MIGRATION INFLOWS ON URBAN FIRMS—PRODUCTION RESTRUCTURING.

Change in product	Any (1)	High education (2)	Low education (3)	High K/L (4)	Low K/L (5)
Panel A: Using unique descriptions to define product categories					
Migration	0.205	-0.001	0.206	0.049	0.157
	(0.049)	(0.028)	(0.036)	(0.038)	(0.026)
Observations	27,062	27,062	27,062	27,062	27,062
Panel B: Using HS codes to define product categories					
Migration	0.120	0.001	0.118	-0.011	0.131
	(0.035)	(0.022)	(0.030)	(0.024)	(0.029)
Observations	27,062	27,062	27,062	27,062	27,062

Key Takeaway: Firms facing a migration induced labor supply shock adjust production lines toward low-skilled, labor-intensive production.

Production Response to Migration (Cont.)

TABLE 8—IMPACT OF MIGRATION INFLOWS ON URBAN FIRMS—TECHNOLOGICAL INNOVATIONS.

New patent	Any (1)	Design (2)	Innovation (3)	Utility (4)
Panel A: Patent categories				
Migration	-0.047	-0.006	-0.035	-0.027
	(0.016)	(0.010)	(0.012)	(0.015)
Observations	31,886	31,886	31,886	31,886
New patent	High education (1)	Low education (2)	High K/L (3)	Low K/L (4)
Panel B: Patent characteristics				
Migration	-0.045	-0.018	-0.046	-0.014
	(0.017)	(0.012)	(0.014)	(0.012)
Observations	31,886	31,886	31,886	31,886

Key Takeaway: Firms in areas with more migrants are less likely to innovate or push the technological frontier.

Conclusion

- Identify causal effect of rural-urban migration on manufacturing production in China
- Instrument migration flows with a shift-share design, relying on exogenous agricultural income shocks
- Migration drives manufacturing firms to become more labor intensive and innovate less
- Results are an important insight for other developing countries

Firm Size

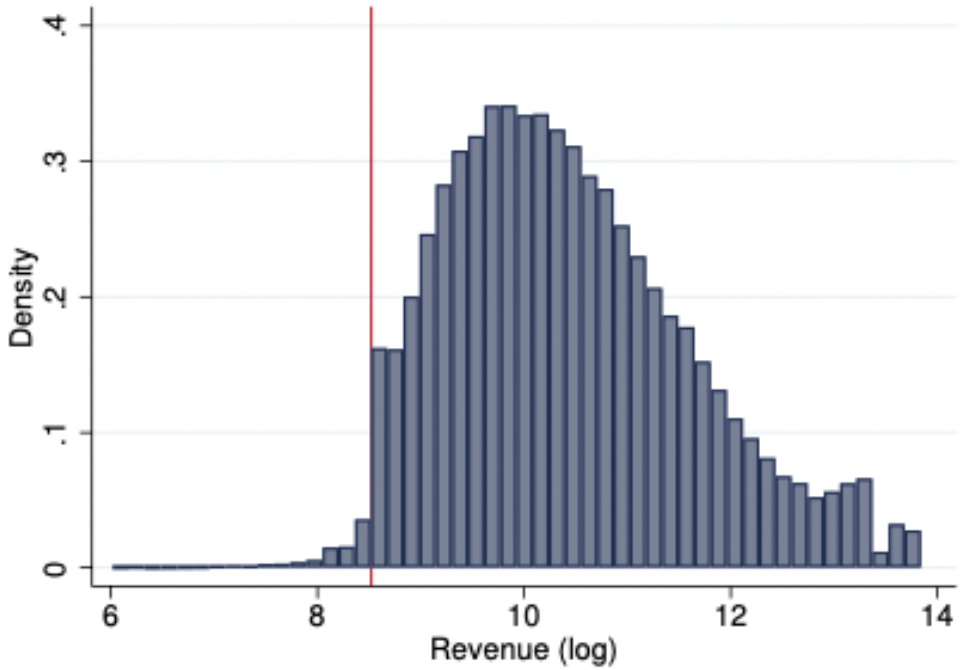
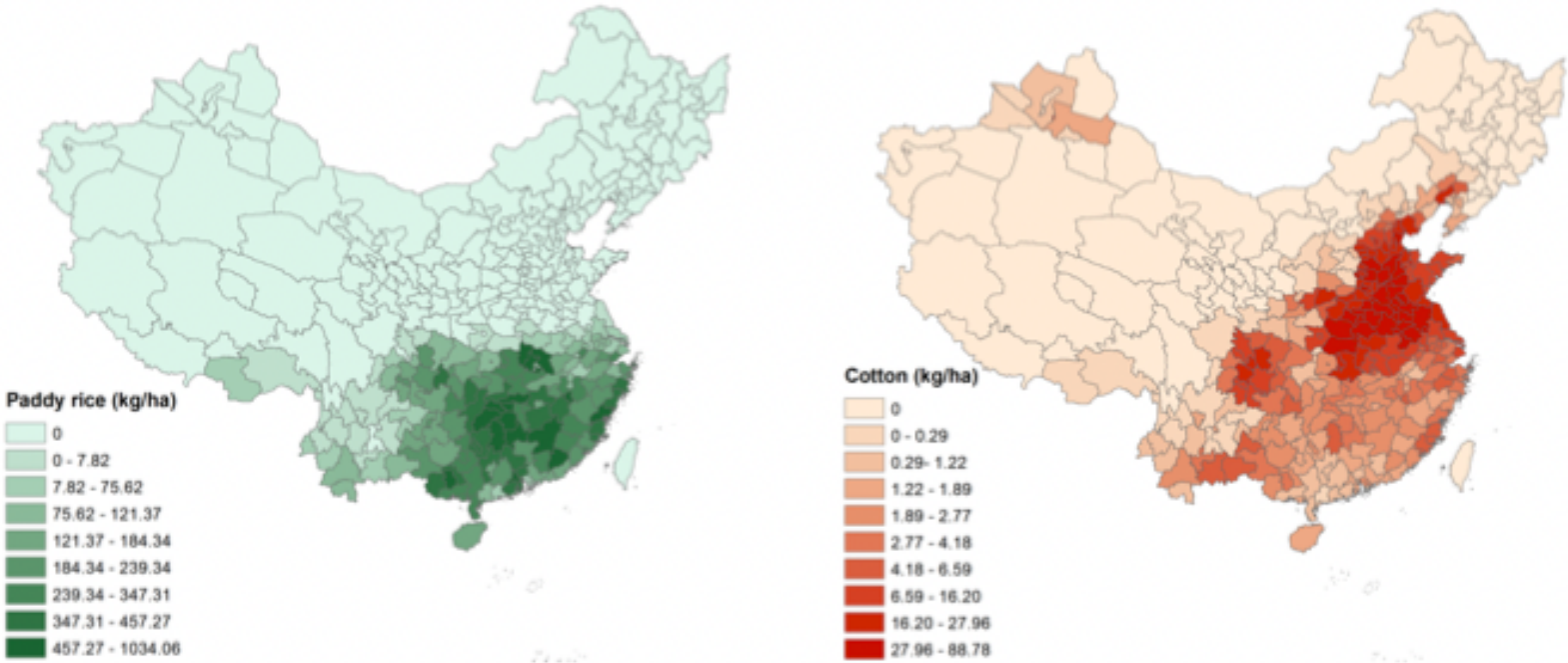


FIGURE 1. DISTRIBUTION OF REVENUE ACROSS FIRMS (NBS, 2000–2006).

Sources: Firm-level data from the National Bureau of Statistics (NBS), 2000–2006. The revenue threshold for appearing in the NBS Census of above-scale firms is RMB 5,000,000, corresponding to $ln(5,000) \approx 8.52$ along the logarithmic scale (of revenues expressed in thousands of RMB).

back

Potential Yield, q_{co}



(a) Paddy rice. (b) Cotton.

FIGURE 2. POTENTIAL YIELD IN CHINA FOR RICE AND COTTON (2000).

Notes: These maps represent the potential yield (GAEZ model) for two common crops in China, i.e., paddy rice (left panel) and cotton (right panel).

back

Robustness I

Concern: agricultural shocks may affect urban production through the intermediate goods market

TABLE 5—IMPACT OF MIGRATION INFLOWS ON URBAN FIRMS—SENSITIVITY ANALYSIS.

	Labor cost (1)	Employment (2)	K/L ratio (3)	Y/L ratio (4)
Panel A: Excluding industries that process agricultural goods				
Migration	-0.139	0.276	-0.408	-0.428
	(0.064)	(0.050)	(0.093)	(0.105)
Observations	29,047	29,047	29,047	29,047
Panel B: Controlling for industry fixed effects				
Migration	-0.104	0.274	-0.472	-0.365
	(0.074)	(0.056)	(0.109)	(0.113)
Observations	31,886	31,886	31,886	31,886

Robustness II

Concern: agricultural shocks may affect rural demand for manufactured goods

Panel C: Excluding migration within a 300-km radius				
Migration	-0.172	0.345	-0.506	-0.513
	(0.071)	(0.065)	(0.115)	(0.129)
Observations	31,886	31,886	31,886	31,886
Panel D: Controlling for shocks in neighboring prefectures				
Migration	-0.241	0.276	-0.838	-0.437
	(0.267)	(0.186)	(0.467)	(0.368)
Observations	31,886	31,886	31,886	31,886
Panel E: Controlling for market access				
Migration	-0.167	0.297	-0.460	-0.440
	(0.066)	(0.053)	(0.102)	(0.112)
Observations	31,886	31,886	31,886	31,886
Panel F: Keeping only exporting firms				
Migration	-0.132	0.257	-0.388	-0.325
	(0.061)	(0.057)	(0.104)	(0.094)
Observations	10,653	10,653	10,653	10,653

Robustness III

Concern: estimates may conflate the effect of current migration with the lagged effect of past migration, which may be endogenous

Panel G: Controlling for the stock of immigrants at baseline				
Migration	-0.110	0.257	-0.559	-0.497
	(0.088)	(0.072)	(0.144)	(0.147)
Observations	31,886	31,886	31,886	31,886

Notes: Standard errors are clustered at the prefecture level and reported between parentheses. The sample is composed of the firms present every year in the NBS firm census between 2000 and 2006. See Section I and Equation (5) for a description of the IV specification.

back