Migrants and Firm's Monopsony Power in the Labor Market *

Lan Lin
Fuzhou University

Qunfeng Wu
Peking University

Mingzhi (Jimmy) Xu
Peking University

January 15, 2026

Abstract

Beyond conventional direct effects, migration expansion reshapes labor market dynamics by altering firms' internal labor composition and the aggregate elasticity of labor supply, thereby recalibrating firms' monopsony power. We develop a general equilibrium model featuring endogenous migrant labor supply and heterogeneous production technologies with task-specific skill requirements. The model explains how increased migrant inflows amplify firms' wage markdowns by altering their labor composition and the aggregate labor supply elasticities they face. Such transformations shift the strategic balance of monopsony power, potentially weakening workers' leverage relative to firms. Exploiting exogenous variation in agricultural income via a Bartik instrument under the Chinese context, we find that firms in cities experiencing larger migrant inflows exhibit a significant rise in wage markdowns. Our mechanism analysis reveals that low-skilled workers face lower labor supply elasticities than high-skilled workers, and consequently, firms with a higher proportion of low-skilled labor operate with lower aggregate supply elasticities and impose greater markdowns. The influx of low-skilled migrants amplifies this distortion by increasing firms' relative employment share of low-skilled workers, thereby further depressing aggregate elasticity and increasing wage markdowns.

Keywords: migrants, markdown, labor composition, monopsony power
JEL Code: D24, J24, J31 , J42

*Lan Lin: School of Economics and Management, Fuzhou University, e-mail: linl2021@econ.pku.edu.cn; Qunfeng Wu: Peking University, e-mail: qfwu@pku.edu.cn; Mingzhi Xu: Peking University, email: mingzhixu@nsd.pku.edu.cn. We are grateful for the financial support of the National Natural Science Foundation of China (Grant No. 72203006, 72322007, 72503230), and the Research Seed Fund of the School of Economics, Peking University. All authors contributed equally and are listed in alphabetical order. All errors are our own.

"Owing to the constantly increasing concentration of leaseholds, Ireland constantly sends her own surplus to the English labour market, and thus forces down wages and lowers the material and moral position of the English working class."

—Karl Marx (1870) *to Sigfrid Meyer and August Vogt*

1 Introduction

The typical trajectory of economic development induces a massive shift of the rural workforce into manufacturing (Lewis, 1954). This influx creates an elastic labor supply that underpins the manufacturing sector's comparative advantage, particularly in labor-intensive industries (Garriga et al., 2023; Liu and Ma, 2023). A growing literature shows that migrants shocks matter for various economic performances: they reshape local economic activity, reallocate labor and capital, and generate heterogeneous effects across firms. Yet, there remains a significant gap in our understanding of how this process reconfigures the relative monopsony power of firms in the labor market. While the intuition that migrant inflows depress wages for comparable native workers is straightforward, a comprehensive evaluation requires a rigorous investigation into the roles of skill-based heterogeneity and the differential elasticities of labor supply across various worker-firm matches.

Wage markdowns have become a central tool for understanding firms' monopsony power in the labor market. By evaluating the wedge between the marginal revenue product of labor and the prevailing wage, markdowns offer an operational measure of monopsony power, revealing how economic rents are partitioned between firms and workers. While firm-level wage markdowns serve as a reliable proxy for monopsony power and maintain a one-to-one mapping with labor supply elasticities, they represent a composite measure of a firm's relationships with workers of varying skill types. Neglecting these granular dynamics may lead to a distorted understanding of firms' wage-setting behavior and yield misguided policy recommendations regarding the welfare of workers with heterogeneous skill profiles.

This paper is the first to estimate the causal effect of rural migrant inflows on urban firm's monopsony power in the labor market. We answer this question by comprehensively evaluate a rigorous investigation into the roles of skill-based heterogeneity and their differential elasticities of labor supply, both theoretically and empirically. We first build a theoretical model, in which rural farmers self-selected to be a migrant workers by utility maximization, which forms endogenous labor supply. Manufacturing firms utilize a combination of tasks with varying skill intensities. A pivotal assumption underlying our theoretical predictions is that low-skilled workers exhibit a more inelastic labor supply than their high-skilled counterparts. This premise is well-documented in the labor economics literature and aligns

with established empirical regularities. Therefore, an exogenous migrants influx shock which alters the internal labor-mix would depresses the aggregate elasticity facing the firm. Consequently, firms exercise greater wage-setting power, leading to higher markdowns, especially within sectors predominantly employing low-skill intensive technologies.

We test our theoretical propositions by exploiting exogenous variation in agricultural income shocks within a Bartik-style instrumental variable framework. Our identification strategy follows the approach developed by Imbert et al. (2022). We instrument migrant flows to Chinese cities using exogenous shocks to agricultural income in rural areas, which trigger rural-urban migration that matches the tremendous rural-urban migration in China during this time window. We use firm census data as well as longitudinal microdata on Chinese manufacturing firms and a population micro-census to measure rural-urban migration.

Our empirical results indicate that firms locate in cities experiencing larger migrant inflows exhibit a significant rise in wage markdowns. The estimates are robust to correction of our Bartik instrument through either the shift measures or the share measures, censoring outliers, various definition of migrant flows. The relationship disappears in placebo exercises that randomly reshuffle migration supply shocks across cities. In our preferred specification, a 10-percentage-point increase in migrants raises firms' markdown between 2000 and 2006 by approximately 0.026 units. Comparing prefectures at the 75th (0.096) and 25th (0.025) percentiles of the actual immigration rate implies a difference in markdown growth of about 0.019 units. This corresponds to an increase of about 2.9 percent. Moreover, heterogeneity analysis supports the baseline results: the capacity for wage suppression is more pronounced among non-SOEs, exporters, and firms in industries characterized by lower capital intensity and higher product market concentration.

We conduct a series of tests to verify the labor compositional channel proposed by the theoretical model. We first show that firms with a higher share of low-skilled workers exhibits greater monopsony power, as evidenced by higher wage markdowns. This pattern is robust across using various kinds of skill intensity definitions using the employment shares of distinct cohorts: (i) those with at most a high school diploma, (ii) associate degree holders, (iii) bachelor's degree holders, and (iv) those with post-graduate qualifications. More importantly, an influx of low-skilled migrants amplifies this pattern, as predicted by our theoretical model. Firms operating in more low-skill–intensive environments benefit more from migration-induced labor supply expansion, thereby exhibiting stronger monopsony power.

Second, rising migrant inflows trigger a compositional shift in firm-level employment, characterized by a declining share of skilled labor. Migrants inflow significantly reduces the average wages of low-skilled resident workers, whereas the estimated effect on high-skilled residents' wages is statistically insignificant and positive. Firms respond to these labor cost

dynamics by altering their skill composition, opting to replace high-skilled personnel with more cost-effective low-skilled migrants. Consequently, we find that cities exposed to larger migrant shocks experience a disproportionate expansion in low-skilled employment relative to high-skilled labor, thereby triggering a fundamental reconfiguration of firms' internal employment structures.

Finally, and most crucially, these compositional shifts translate into significant changes in firms' aggregate labor supply elasticities, which ultimately govern the determination of wage markdowns. We estimate skill–specific labor supply elasticities follow the mothodology proposed by Khanna et al. (2025). We provide empirical evidence that low-skilled labor possesses a significantly more inelastic supply curve compared to the more elastic supply of high-skilled counterparts. As such, the expansion of the low-skilled migrant pool depresses the firm-level aggregate labor supply elasticity through a compositional channel. Collectively, these results offer compelling support for the theoretical propositions of our model.

Related Literature Our paper contributes to four distinct strands of literature. First, a broad literature has examined the effects of migration on a wide range of economic outcomes, including economic growth, productivity, innovation, and fiscal performance (Angelucci, 2015; Bryan and Morten, 2019; Tombe and Zhu, 2019; Imbert and Papp, 2020). Within this broader context, our paper relates to the literature examining the impact of migration on labor market outcomes, particularly wages and employment, which has documented substantially divergent conclusions across studies (Behrens and Sato, 2011; Brücker and Jahn, 2011; Accetturo et al., 2012; Kerr et al., 2015). Evidence from large and sudden migration shocks suggests that migrant inflows need not generate sizable declines in native wages or employment, underscoring the role of general equilibrium adjustments and local absorption mechanisms (Card, 1990). In contrast, more pronounced negative wage effects arise when labor markets are defined at finer skill levels, emphasizing education–experience–specific labor supply competition (Borjas, 2003). In addition, these differences can be partly reconciled once factor substitutability and capital adjustment are taken into account, with imperfect substitution between native and migrant workers and endogenous capital responses substantially attenuating competitive wage pressures (Ottaviano and Peri, 2012). Moving beyond average effects, migration has been shown to affect the wage distribution heterogeneously, with adverse impacts concentrated at the lower tail, highlighting the importance of distributional analysis in assessing migration-induced labor supply shocks (Dustmann et al., 2013).

Meanwhile, a growing literature studies how migration affects firms' labor market behavior at destination. Changes in low-skilled labor supply have been shown to influence firms'

adoption of automation and capital–skill complementarity (Lewis, 2011), while firms also respond to migration shocks through adjustments in production processes and product mix, adopting technologies biased toward more abundant labor inputs (Gray et al., 2020; Monras, 2020). In contrast with a literature that focuses on international migration (Grossmann and Stadelmann, 2013), recent work emphasizes internal migration in developing economies, where large-scale labor reallocation is a central feature of structural transformation (Imbert et al., 2022). Importantly, this approach exploits origin-area agricultural income shocks to instrument migrant inflows, providing a credible strategy for isolating exogenous labor supply variation. Building on this identification framework, this paper studies how migration-induced labor supply shocks affect firms' labor market power, thereby linking firm-level production adjustments to distributional outcomes in the labor market.

Second, our focus on migration as a source of large labor supply shocks contributes to the literature on the determination of wage markdowns. A growing body of work examines wage determination in labor markets characterized by imperfect competition, where firms possess wage-setting power and pay wages below workers' marginal revenue product. In such environments, wage markdowns arise from search frictions, mobility costs, and limited outside options faced by workers, implying that wages reflect not only productivity but also firms' labor market power (Mortensen and Burdett, 1998; Manning, 2003a, 2011; Card et al., 2018). Recent research emphasizes that wage markdowns also respond to policy reforms and external shocks that alter labor market conditions. In particular, evidence from China shows that trade liberalization weakens firms' labor market power, reduces the dispersion of wage markdowns across firms, and improves allocative efficiency, with intermediate input tariff reductions playing a particularly important role and effects that are stronger in skill-intensive firms (Pham, 2023; Xie et al., 2024; Kondo et al., 2024). While these studies exploit trade reforms using Chinese data, it is important to note that the same period also witnessed large-scale internal migration, suggesting that labor supply shocks may constitute an additional and independent force shaping firms' labor market power. Motivated by this observation, this paper studies internal migration in a developing-country context. The most closely related contribution is Amior and Manning (2025) , who embed migration in a monopsonistic wage-setting framework and show that migration can affect wages through changes in firms' wage markdowns rather than solely through marginal productivity. Using U.S. data, they demonstrate that labor market power plays a central role in shaping wage responses to migration. In our context, we focus on internal migration within an emerging economy and identifies a distinct compositional channel through which migration reshapes firms' labor monopsony. We provide empirical evidence that low-skilled labor exhibits a significantly more inelastic labor supply than high-skilled labor. As a result, the expansion of the low-

skilled migrant workforce reduces firms' aggregate labor supply elasticity through changes in labor composition, thereby strengthening firms' labor market power and amplifying wage markdowns.

Our study is also closely related to the literature on firms' skill composition and the organization of tasks within production. This literature emphasizes that firms combine tasks with different skill intensities, and that shocks to labor supply can induce endogenous adjustments in task allocation and workforce composition (Autor and Murnane, 2003). A broad literature examines how trade, technological change, and organizational restructuring shape firms' skill composition through adjustments in production organization and task allocation (Grossman and Rossi-Hansberg, 2008; Acemoglu and Autor, 2011). A central strand highlights occupational sorting and task reallocation as key mechanisms through which labor markets absorb migration shocks (Ottaviano et al., 2013). Migrants tend to specialize in manual- and routine-intensive tasks, allowing native workers to reallocate toward communication- and cognitive-intensive occupations and thereby mitigating direct wage competition (Sparber and Peri, 2009). Native workers are shown to respond to migrant inflows by upgrading their occupational task content, leading to limited wage losses and, in some cases, improved employment outcomes (Foged and Peri, 2016). Complementing these mechanisms, early theoretical work shows that trade liberalization can reshape the skill composition of the labor force by altering migration incentives across skill groups, increasing the migration of unskilled workers while leaving skilled migration largely unaffected in developing economies, thereby influencing the overall skill mix available to firms (Lopez and Schiff, 1998). Building on these insights, this paper studies a setting in which migration is predominantly low-skilled and examines how such inflows affect firms' wage-setting behavior through changes in skill composition. We develop a theoretical framework in which manufacturing firms rely on a bundle of tasks with different skill intensities, and in which low-skilled labor is characterized by a more inelastic labor supply than high-skilled labor. In this environment, an exogenous migration shock that expands the pool of low-skilled labor alters the internal task and skill mix within firms and depresses the aggregate labor supply elasticity they face. This compositional channel strengthens firms' labor market power and provides a mechanism through which migration can affect wage outcomes beyond standard competitive channels.

Our paper finally relates to the literature studying the relationship between the agricultural sector and manufacturing activity in the process of structural transformation, with particular emphasis on how agricultural conditions affect industrial development and urban labor markets (Lewis, 1954; Gollin et al., 2002; Herrendorf et al., 2013). In this view, adverse agricultural shocks increase rural–urban migration and expand the supply of labor to urban manufacturing, while positive shocks can slow migration and tighten urban la-

bor markets (Groeger and Zylberberg, 2016; Minale, 2018). These mechanisms highlight agriculture as an important "labor push" factor in the process of development. A related empirical literature exploits exogenous variation in agricultural income, arising from weather shocks, rainfall variability, crop price fluctuations, or agricultural policy reforms, to identify the causal effects of rural conditions on migration and urban labor markets (Kaivan, 2003; Feng et al., 2017; Kleemans and Magruder, 2018). While much of this literature focuses on aggregate employment and wage outcomes under competitive labor market assumptions, related evidence shows that agricultural shocks can have persistent effects on the structure of manufacturing by inducing specialization in low-productivity activities through the inflow of unskilled labor released from agriculture (Bustos et al., 2019).

Building on this literature, our empirical strategy exploits exogenous variation in agricultural income as an instrument for migration-driven labor supply shocks and studies their effects on urban manufacturing firms from a labor market power perspective. By using agricultural income shocks, we identify migration flows that are driven by conditions in the rural sector rather than by changes in labor demand in urban areas, distinguishing our approach from "labor pull" mechanisms in which migration responds endogenously to productivity growth in manufacturing and cities (Facchini et al., 2019). Our contribution is to show that agricultural shocks affect urban manufacturing not only by changing the quantity of labor supplied, but also by altering firms' labor supply elasticity and wage-setting power. By focusing on firm-level outcomes, this paper provides new evidence on how shocks originating in agriculture shape urban labor markets during structural transformation through firms' labor monopsony.

The rest of the paper is organized as follows. Section 2 outlines the institutional context of China's rural economy, migration patterns, and firm's monopsony patterns. Section 3 introduces the a general equilibrium model explaining how migrants affect firm's markdown through labor compositional channel. Section 4 describes the data and presents the empirical analysis. Section 5 provides baseline empirical results and various robustness checks and heterogeneity analysis. Section 6 test mechanisms that fit the model assumptions and predictions. Section 7 concludes.

2 Institutional Background

The Great Migration in China Historically, China's economy was predominantly agrarian, with agriculture accounting for the majority of both population employment and economic activity. Back to the early reform years, over eighty percent of its population resides in the rural area. This share has declined gradually since then, accompanied with China's

opening up to the world market and domestic institutional reform, e.g., its Hukou system change. By the end of 2006, China's rural population share declined to 56.1%. This corresponds to over 200 million rural-to-urban migrants since 1980s [1], which is more than four times the volume of the Great Migration of Europeans to North America from 1800 to World War I (Facchini et al., 2019).

Two main forces were driving this dramatic process. On the one hand, low-income rural residents desired outside working opportunities to increase their income through migration. Like many other developing countries, China's agricultural sector was full of vulnerability, frequently haunted by unpredictable outside shocks, e.g., poor weather condition, inadequate market access, volatile output prices, and so on. Crowded with billions of farmers in the rural area by then, China's agriculture sector exhibits typical labor-intensive traditions, hardly embodied with any modern infrastructures or frontier techniques (Huang et al., 2008). Yearly rural per capita disposable income was only 2282 RMB yuan (less than 300 US dollars) in 2000, which is almost just one third that of the urban residents (6256 RMB yuan).

On the other hand, China's opening to the global market has created substantial economic opportunities for the urban manufacturing sector. Therefore, for migrants from the agricultural sector, especially after entering the WTO in 2001. Increasing outside world product demand motivated Chinese firms to expand their production and as a result, increasing labor employment which enlarge labor demand. As Feenstra and Hong (2007) pointed out, export growth in China contributed 2.5 million jobs per year during 1997-2002. Urban manufacturing sector provides much higher returns compare with farming in the rural area, inducing more rural residents to migrate.

The unprecedented scale of rural-to-urban migration has substantially altered urban labor markets by sharply increasing the supply of low-skilled labor (Ge and Yang, 2014; Imbert et al., 2022). Yet the existing literature largely overlooks its implications for firms' bargaining power in the labor market. We aim to fill this gap by empirically testing the effect of the increase in migrant labor during this period on firms' wage markdowns.

Firm's Market Power and Labor Composition Firms typically possess market power on their input markets, i.e., labor market or intermediate input market (Rubens, 2023). Understanding firm's market power on the labor market is essential for quantifying firm's buyer power, up-stream and down-stream allocative inefficiency, as well as capital-labor income distribution (Yeh et al., 2022). As shown in Figure A.1, during the period from 2002 to 2006, Chinese firms' wage markdowns have gradually increased, indicating a growing

[1] Specifically, Imbert et al. (2022) estimate that 45 million rural workers migrated to cities between 2000 and 2005, equivalent to 16 percent of the urban population in 2000.

monopsony power in the labor market. In contrast, the wage ratio between low-skilled and high-skilled workers has continuously decreased over the same period, reflecting a widening wage gap between worker groups with different skill intensities.

An exogenous large increase in migrants inflow could contribute to this pattern by shifting firms' labor composition and aggregate labor supply elasticities. Rural-to-urban migrants are typically characterized as low-skilled workers in urban manufacturing sectors. Consequently, an exogenous expansion in low-skilled labor supply would reduce skill intensity across urban manufacturing firms and industries. Naturally, it not only lowers wages of low-skilled workers, but also enlarges the skill-premium between workers.

As shown in Figure A.2, firms that with more high-skilled workers, both in level and in share, exhibit lower markdowns, indicating weaker monopsony power in local labor markets. Consequently, shifts in migrant inflows impact firms' wage markdowns through a compositional channel—a mechanism that has been largely overlooked in the existing literature.

3 A Model of Migrants' Effect on Firm's Markdown

This section develops a general equilibrium theoretical framework to analyze the impact of an exogenous increase in rural-to-urban migration on the labor composition and labor markdown of urban manufacturing firms, thereby providing a theoretical foundation for the empirical analysis in subsequent sections.

For analytical tractability, the model posits an economy comprising a single rural region and a single urban region. In the rural sector, agents are primarily engaged in agricultural production. Migration is modeled as an individual optimization decision: agents choose to relocate to the urban manufacturing sector to become migrant workers and earn corresponding wages by comparing their expected returns across the two sectors.(Morten, 2019; Gai et al., 2021). Negative income shocks in the agricultural sector, such as adverse commodity price movements or severe weather events, decrease the agricultural opportunity cost of migration, thereby stimulating out-migration from rural areas.

In the urban region, the economy is characterized by S distinct manufacturing sectors, indexed $s = 1, 2, \ldots, S$. Each sector s consists of a set of N_s firms that produce differentiated output by utilizing various labor types to perform a range of production tasks. These urban manufacturing firms hire three categories of labor input: urban high-skill workers, urban low-skill workers, and rural-to-urban migrant workers. To produce output within each sector, firms employ low-skill and high-skill workers to execute a continuum of tasks. Crucially, we initially posit that tasks assigned to urban low-skill workers can also be performed by migrant workers, characterized by a constant elasticity of substitution (CES) σ. However,

tasks performed by high-skill workers are assumed to be non-substitutable with either low-skill or migrant labor. There are a fixed number of $\bar{L}^r$ rural residents, who can choose to migrate to urban sectors or not according to their utility maximization decision. For simplicity and also to focus on our main purpose in this paper, we assume that urban residents are immobile, with total labor supply of low-skill and hign-skill workers fixed at $\bar{L}^l$ and $\bar{L}^h$, respectively.

3.1 Migration Labor Supply

A farmer i from the rural region, who chooses to migrate to and work in the urban sector s, firm f, draws her indirect utility from following function:

$$V_{isf} = ln(w^*_{sf}) - ln(w^r) - ln(d) + \epsilon_{isf} \tag{1}$$

where w^*_{sf} denotes wages she earns working in firm f of the urban manufacturing sector s, w^r denotes how much she can earn in the rural region, i.e., from farming through the rural agriculture sector. Lower (higher) farming income would increase (reduce) her willingness to migration since it reduces (increases) her opportunity cost to migrate. d refers to migration cost, i.e., transportation cost and so on. The error term, ϵ_{isf}, is assumed to be distributed independently across farmers according to a generalized extreme value distribution:

$$G_\varepsilon(\varepsilon) = \exp\left\{ - \left[\sum_{s \in S} \left(\sum_{f \in N_s} \exp\left(-\frac{\varepsilon_{isf}}{\kappa_s} \right) \right)^{\frac{\kappa_s}{\kappa}} \right]^{\kappa} \right\}, \tag{2}$$

where S is the set of industries, N_s is the set of firms in industry s, $0 < \kappa_s < 1$ is the within-industry (nest) scale parameter, and $\kappa > 0$ is the across-industry (industry-layer) scale parameter.

Given the above setting, the labor supply of migrants is:[2]

$$q_{sf} = (\bar{L}^r \lambda_s) e^{\frac{1}{\kappa_s} \ln \frac{w^*_{sf}}{w^r d}} \tag{3}$$

where λ_s is a constant. A reduction in migration costs (d) or rural income (w^r) induces a larger flow of rural workers into the urban sector. Correspondingly, the labor supply elasticity of migrant workers is given by $\frac{1}{\kappa_s}$, implying that the firm's wage markdown for this group is $1 + \kappa_s$.

[2] Appendix C provides a formal proof of the derivation.

3.2 Production and Labor Composition in the Urban Manufacturing Sectors

Production Function and Task Assignments In each manufacturing sector, there is a representative firm producing following Cobb-Douglas functional form:

$$Y_s = A_s l_s^{\theta_{sl}} h_s^{1-\theta_{sl}} \tag{4}$$

where A_s is a technological parameter, l_s and h_s represent low-skill and high-skill worker composite tasks, respectively. θ_{sl} is the cost-share of the low-skill factor of sector s. While high-skill worker composite tasks are assumed to be performed only by urban high-skill workers, low-skill tasks can be performed by both urban local residents and migrants from the rural region, given by:

$$l_s = \left[\int_0^1 l_s^{\frac{\sigma-1}{\sigma}}(j) dj \right]^{\frac{\sigma}{\sigma-1}} \tag{5}$$

where $l_s(j)$ is the input of task j in industry s and $\sigma > 0$ is the elasticity of substitution between tasks. A unit of task output is produced using one unit of low-skill labor when a task is performed for industry s. Furthermore, low-skill tasks are assumed to be ordered by decreasing degree of operability, defined as a unit cost parameter $\tau(j)$ for any $j \in [0, 1]$, by migrants from the rural area, that is:

$$\frac{\partial \tau(j)}{\partial j} > 0 \tag{6}$$

Denote wages for urban low-skill residents and for rural migrants as w and w^* [3], respectively. Then, the unit price of any particular low-skill task j, denoted as $c_{sl}(j)$, is assumed to be equal to its unit cost, can be written as:

$$c_{sl}(j) = \begin{cases} w^* \tau(j) & \text{if } 0 \leq j < I_M \\ w & \text{if } I_M < j \leq 1 \end{cases} \tag{7}$$

where I_M is the marginal task that is indifferent in being produced by urban local low-skill residents or by migrants from the rural region, which can be obtained by the following equilibrium condition:

$$w^* \tau(I_M) = w \tag{8}$$

[3] Wages are determined at the firm level within each sector. For the sake of clarity, we suppress firm and sector indices throughout this section.

Holding urban low-skill workers' wage w fixed, an exogenous increase in the labor supply of migrant workers would lower their wages w^*, and expand the task range that migrant workers can take according to equation (8).

Demand for Production Tasks The demand for low-skill labor j in industry s, $l_s(j)$, and for high-skill labor h_s, is given by:

$$l_s(j) = \frac{Y_s}{A_s}\left[\frac{c_{sl}(j)}{P_{sl}}\right]^{-\sigma}\left[\frac{\theta_{sl}}{1-\theta_{sl}}\frac{P_{sl}}{P_h}\right]^{1-\theta_{sl}} \tag{9}$$

$$h_s = \frac{Y_s}{A_s}\left(\frac{P_{sl}}{P_h}\frac{1-\theta_{sl}}{\theta_{sl}}\right)^{\theta_{sl}} \tag{10}$$

where P_h is the wage for the high-skill composite tasks, and P_{sl} is the exact price index for the low-skill tasks composite, given by:

$$P_{sl} = \left[\int_0^1 c_{sl}^{1-\sigma}(j)dj\right]^{\frac{1}{1-\sigma}} = \left[\int_0^{I_M}(w^*\tau(j))^{1-\sigma}dj + (1-I_M)w^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{11}$$

Define

$$\Omega(I_M) \equiv \left\{\int_0^{I_M}\left[\frac{\tau(j)}{\tau(I_M)}\right]^{1-\sigma}dj + (1-I_M)\right\}^{\frac{1}{1-\sigma}}$$

then $\Omega'(I_M) < 0$, and $P_{sl} \equiv w\Omega(I_M)$.

Denote the migration labor needed to perform low-skill task j in each industry s as $D_{sl}(j) = l_s(j)$. Their aggregate across all tasks as D_{sl} consists of tow parts: $D_{sl}^M = \int_0^{I_M} D_{sl}(j)dj$, and total demand for urban low-skill labor as $D_{sl}^U = \int_{I_M}^1 D_{sl}(j)dj$, then we would have:

$$D_{sl}^M = \int_0^{I_M} D_{sl}(j)dj = B_s Y_s\left(\frac{P_{sl}}{P_h}\right)^{1-\theta_{sl}}\int_0^{I_M}\left(\frac{w^*\tau(j)}{P_{sl}}\right)^{-\sigma}dj \tag{12}$$

$$D_{sl}^U = \int_{I_M}^1 D_{sl}(j)dj = B_s Y_s\left(\frac{P_{sl}}{P_h}\right)^{1-\theta_{sl}}(1-I_M)\left(\frac{P_{sl}}{w}\right)^{\sigma} \tag{13}$$

where $B_s = \frac{1}{A_s}\left(\frac{\theta_{sl}}{1-\theta_{sl}}\right)^{1-\theta_{sl}}$.

Labor Market Equilibrium In equilibrium, migrant workers' labor supply equals to their labor demand, that is, $q^s = N_s D_{sl}^M$ for all $s \in S$. Therefore, and exogenous increase in migrant workers' labor supply q^s would lower their wages, i.e., w^*, according to this labor market clearing condition. Holding other variables constant, it would increase migrant

workers' task range, i.e., I_M, according to equation (8), and then results in an increase in firm's aggregate markdown following process we introduce below in the next section.

3.3 the Effect of Migrants on Firm's Markdown

We assume the labor supply of both the high-skill workers and low-skill urban workers are fixed at $\bar{L}^h$ and $\bar{L}^l$, respectively. Consequently, the firm's markdown for these two specific labor types is equal to unity.[4] However, migrant workers face an upward shaped labor supply function with an elasticity of $1/\kappa_s$ for firms each industry s. Therefore, firm's markdown over migrant workers equals to $1+\kappa_s$, where κ_s is the inverse labor supply elasticity of migrant workers according to equation (3). Therefore, given the structure of the labor market, the markdown for firms in each sector s can be written as:

$$\varphi_s = \left(\frac{\left[\int_0^{I_M} \left(w^* \left(1+\kappa_s\right) \tau(j) \right)^{1-\sigma} dj + \int_{I_M}^1 \left(w\right)^{1-\sigma} dj \right]^{\frac{1}{1-\sigma}}}{P_{sl}} \right)^{\theta_{sl}}, \tag{14}$$

which can be further written as a function of κ_s and I_M, as well as parameters σ and θ_{sl}:

$$\varphi_s = \left(\frac{\left((1+\kappa_s)^{1-\sigma} \ \left[\Omega(I_M)^{1-\sigma} - (1-I_M) \right] + (1-I_M) \right)^{\frac{1}{1-\sigma}}}{\Omega(I_M)} \right)^{\theta_{sl}}, \tag{15}$$

Equation (15) has several properties that are vital for our understanding over firm's monopsony power in the labor market. First, Lower labor supply elasticity of migrant workers than high-skill types of labor, i.e., $1/\kappa_s < 1$, drives the effect of changes in the relative supply of migrant workers on firm's markdown. Lower labor supply elasticities allows for higher level of exploitations, thus higher markdowns.

Secondly, firms in industries with higher skill intensity, i.e., lower value of θ_{sl}, would have lower markdown in the labor market. This difference is the result of the compositional effect of low-skill tasks and high-skill tasks over different skill intensities.

Thirdly and most importantly, since $\Omega(I_M)$ decreases with I_M, it can be easily proved that

$$\frac{\partial \varphi_s}{\partial I_M} > 0 \tag{16}$$

[4] We adopt this assumption for ease of exposition. The foundational assumption underpinning our theoretical predictions is that migrant workers face a lower labor supply elasticity than high-skilled workers (specifically, $1/\kappa_s < 1$). We formally test this premise in our empirical analysis, and find it to be consistent with the established literature (Xie et al., 2024; Khanna et al., 2025).

Further more, we can also have:

$$\frac{\partial^2 \varphi_s}{\partial I_M \partial \theta_{sl}} > 0 \tag{17}$$

Ceteris paribus, an exogenous increase in the migrant labor supply lowers the equilibrium wage (w^*) and, consequently, increases I_M as dictated by equation (8). Therefore, an influx of migrants intensifies firm-level markdowns. This effect is disproportionately stronger in less skill-intensive industries—those characterized by a higher θ_{sl}—where the expansion of the migrant labor supply has a more pronounced impact on market power.

Drawing upon the equilibrium characterization derived above, we formalize the core comparative static results through the following two propositions:

Proposition 1. *An exogenous increase of rural migrants leads to a direct increase in the demand for low-skill tasks, results in an increase in the task cutoff I_M and an increase in firm's wage markdown. At the same time, firms' skill intensity decreases with the increase of migrants.*

Proposition 2. *For firms in industries with higher low-skill technologies, i.e., higher value of θ_{sl}, the positive effect of migrants on firms' markdown would be stronger.*

4 Empirical Framework

This section outlines the empirical framework employed to examine the impact of migration on firms' labor market power, guided by the theoretical predictions developed in the preceding section.To this end, our identification strategy exploits exogenous shocks to agricultural income as a 'push factor' for rural-to-urban migration. Following the approach of Imbert et al. (2022), we leverage these income fluctuations to generate exogenous variation in labor supply across local labor markets. We first detail the comprehensive data sources utilized for our empirical analysis, followed by a discussion of the identification strategy employed to establish causality, which builds upon the framework developed by Imbert et al. (2022).

4.1 Data Description

We first outline the primary data sources used to estimate wage markdowns and to identify the effects of exogenous migrant inflows. Comprehensive details regarding data construction, cleaning procedures, and summary statistics are relegated to Appendix B.1.

Chinese Manufacturing Firms Dataset Our primary firm-level data are drawn from the Annual Survey of Industrial Firms (ASIF), a comprehensive census conducted by China's National Bureau of Statistics (NBS). ASIF is an annual administrative census/survey of above-scale industrial enterprises with nationwide coverage across ownership types (state-owned, collective, private, and foreign-funded) at the legal-entity level. The dataset has been widely used in existing literature on Chinese manufacturing firms (e.g., Brandt et al., 2017; Xie et al., 2024). This dataset covers manufacturing firms with annual sales at or above RMB 5 million and includes each firm's unique identifier and registry information (enterprise code, ownership, registration details, four-digit industry code, province/prefecture location), together with comprehensive financial and employment records, including employment and payroll, revenues and costs, and balance-sheet items such as total assets, total liabilities, equity, and profits.

Firm census data To measure firms' skill intensity and examine changes in labor composition over time, we draw on firm-level data from the China Industrial Census 1995 and the China Economic Census 2004, compiled by the National Bureau of Statistics of China. These two census datasets provide exceptionally comprehensive coverage of China's manufacturing sector, spanning a ten-year period and including industrial firms of all sizes, rather than being restricted to firms above a size threshold (Long and Zhang, 2011). The rich employment information contained in the censuses allows us to construct firm-level measures of labor composition by educational attainment and to track changes in employment structure within firms over time. Meanwhile, both censuses report firm identifiers that enable us to merge the census information with the ASIF database, thereby linking detailed skill composition measures to firms' markdowns that capture labor market power. The broad coverage and consistency of the census data make them particularly well suited for studying how migration-induced labor supply shocks are reflected in firms' skill composition and employment decisions.

Population Census Data As is standard in the literature, bilateral regional-level migration flow information is obtained from the population census data. We use the nationally representative 2005 1% Population Survey ("Mini-Census") administered by the NBS, which enumerates individuals at their current place of residence—migrants included—and records hukou type (agricultural vs. non-agricultural), place of registration and residence (prefecture level), occupation, industry, income, education, and housing. Crucially for flows, the survey asks the main reason for leaving the place of registration, the year of departure, and place of residence one and five years prior to the interview.

Agricultural Data To construct exogenous income shocks for rural migrants from the agriculture sector, we leverage the information provided by the FAO/IIASA Global Agro-Ecological Zones (GAEZ). Baseline cropping portfolios at the prefecture level are measured from a geo-coded map of harvested areas in China from the 2000 World Census of Agriculture, which we combine with potential yields from the FAO/IIASA Global Agro-Ecological Zones (GAEZ). For each prefecture–crop cell, baseline agricultural potential is the product of harvested area and GAEZ potential yield; these portfolios are normalized to shares and held fixed over our sample window. International crop prices are taken from FAO Agricultural Producer Prices (with robustness to the World Bank "Pink Sheet" series). For each major crop, we construct a world price series excluding China to avoid domestic endogeneity. Origin–year agricultural shocks are then formed by interacting the time-invariant prefecture portfolios with the external (world) price movements for the corresponding crops.

4.2 Wage Markdown and Migrants Measurement

Markdown Our primary dependent variable is firm-level labor market power, quantified by the markdown of wages relative to the marginal revenue product of labor (MRPL). This measure captures the extent to which firms exercise monopsony power, a methodology widely adopted in the recent literature(Yeh et al., 2022; Rubens, 2023; Rubens et al., 2024). This wedge manifests as a gap between the marginal revenue product of labor (MRPL) and the observed wage. Consequently, wage markdown is defined as follows:

$$markdown_{it} \equiv \frac{MRPL_{it}}{Wage_{it}} \tag{18}$$

which directly measures the extent to which firms extract labor surplus.[5]

To measure the markdown, we follow the approach proposed by Yeh et al. (2022). Specifically, we first compute firm-level markups μ_{it} using a material input k, as:

$$\mu_{it} = \frac{\theta_{it}^{k}}{\alpha_{it}^{k}}, \tag{19}$$

where θ_{it}^{k} is the output elasticity of material input k and α_{it}^{k} is its share in firm revenue.

Assuming a perfectly competitive market for intermediate goods, along with a differentiable and finitely elastic wage schedule, the markdown, defined as the ratio of the marginal product of labor to the wage, is then calculated as:

$$\text{markdown}_{it} = \frac{\theta_{it}^{\ell}}{\alpha_{it}^{\ell}} \cdot \mu_{it}^{-1}, \tag{20}$$

[5] Appendix B.2 provides more details about the definition of firm level markdown.

where θ_{it}^{ℓ} denotes the output elasticity of labor and α_{it}^{ℓ} is the share of labor's revenue.

We estimate the necessary elasticities using the DLW method (De Loecker and Warzynski, 2012), enabling the construction of firm-level markdowns that capture deviations from competitive wage-setting behavior. More details on the estimation of elasticities and the construction of markups and markdowns are provided in Appendix B.2.

Migrants The key explanatory variable of interest is the migrant inflow intensity, measured as the cumulative immigration rate at the prefecture level between 2000 and 2005. Formally, following Imbert et al. (2022), we first compute the annual immigration rate m_{dt} as the number of rural-to-urban migrants arriving in destination d in year t, divided by the baseline non-migrant population in d in the year 2000, denoted as P_d:

$$m_{dt} = \frac{\sum_{o \in \Theta \setminus \{d\}} M_{odt}}{P_d}, \tag{21}$$

where M_{odt} denotes the number of migrants from origin o to destination d in year t, and Θ represents the set of all prefectures. We then aggregate the yearly immigration rates to obtain the total immigration rate over the period:

$$m_d = \sum_{t=2000}^{2005} m_{dt}. \tag{22}$$

This cumulative measure captures sustained changes in local labor supply due to migration during a period of rapid structural transformation in China. By normalizing flows using the 2000 baseline population, the measure reflects relative increases in labor supply across cities of different sizes and initial demographic conditions. [6]

Figure A.3 displays the immigration rates across different prefectures in China from 2000 to 2005. It highlights the substantial variation in migration flows, with notable concentrations in the eastern coastal regions and some western areas.

4.3 Empirical Strategy and Identification

Empirical Strategy Our baseline specification relies on a balanced panel of firms continuously observed from 2000 to 2006 and examines the relationship between migration inflows and changes in markdown. Specifically, we exploit variation in migration inflows occurring between 2000 and 2005, and relate it to subsequent changes in firm-level markdowns measured between 2000 and 2006. We regress the change in firm's markdown between 2000 and 2006, $\Delta\text{markdown}_{id}$, for a firm i in prefecture d, on the cumulative migration flows to

[6] More details on the data source of migrants variables are provided in Appendix B.1.

prefecture d over the period 2000-2005, denoted by m_d:

$$\Delta \text{markdown}_{id} = \alpha + \beta m_d + \varepsilon_{id} \tag{23}$$

where standard errors are clustered at the prefecture level for the destination. Migration flows may reflect a surge in labor demand, which induces a reversal causality in the OLS estimate. We therefore construct a labor supply shock instrument and estimate equation (23) with two-stage least squares (2SLS).

Instrumental variable construction. To isolate the supply-driven component of migration and address endogeneity from demand-side shocks at destination, we use a Bartik (shift–share) instrument following Imbert et al. (2022). Let d index destination prefectures, $o \in \Theta$ origin prefectures, and c crops.

Shifts (origin agricultural-income shocks). For each origin–year (o, t), construct an agricultural-income shock by interacting baseline cropping portfolios with innovations in world crop prices. Denote baseline harvested area for crop c in origin o by h_{co} (from the geo-coded map of harvested areas in China, 2000 World Census of Agriculture) and potential yield by y_{co} (FAO/IIASA GAEZ). Potential output is

$$q_{co} \equiv h_{co}\, y_{co}\,,$$

held fixed over the sample. Let $\bar{p}_c$ be the baseline international price (excluding China), and let $\hat{\varepsilon}_{ct}$ be the AR(1) innovation in $\log p_{ct}$ (crop–year price residual). The origin–year shock is the portfolio-weighted average price innovation: [7]

$$s_{ot} = \frac{\sum_c \bar{p}_c\, q_{co}\, \hat{\varepsilon}_{ct}}{\sum_c \bar{p}_c\, q_{co}}\,, \qquad s_o \equiv \sum_{t=2000}^{2005} s_{ot}. \tag{24}$$

Shares (historical settlement patterns). Historical migration links from o to d are measured using the stock of rural-hukou migrants arriving before 2000. Let M_{odt} be migrants from o to d in year t; then the pre-2000 stock is $\sum_{t<2000} M_{odt}$ and the bilateral share is

$$\lambda_{od} = \frac{\sum_{t<2000} M_{odt}}{\sum_{d'} \sum_{t<2000} M_{odt}}\,. \tag{25}$$

Bartik instrument at destination. The instrument for migrant inflows to destination d

[7] In the main analysis we use a balanced firm panel for 2000–2006, so we cumulate origin shocks over 2000–2005, i.e., $s_o = \sum_{t=2000}^{2005} s_{ot}$. When examining the impact of migration on skill intensity, we rebuild the instrument by cumulating s_{ot} over 1995–2004 to align with the education data window.

aggregates origin shocks using historical shares:

$$z_d = \sum_{o \in \Theta \setminus \{d\}} \lambda_{od} \, s_o \, . \tag{26}$$

By construction, z_d varies across destinations with the interaction of exogenous origin shocks and pre-determined settlement patterns.

To assess the relevance of our instrument, we estimate the following first-stage regression, in which the actual immigration rate between 2000 and 2005, m_d, is regressed on the instrument z_d:

$$m_d = \beta_0 + \beta_1 z_d + \eta_d. \tag{27}$$

Table A.1 reports the corresponding estimates. We find a statistically significant and negative relationship between z_d and immigration inflows, confirming that the instrument is a strong predictor of immigration and thus satisfies the relevance condition.

5 Empirical Results, Robustness Checks, and Heterogeneous Analysis

This section presents the empirical results of the effect of increasing migrants on firms' markdowns. Our identification strategy is borrowed from Imbert et al. (2022). We explore how changes in local labor supply driven by migration flows affect the extent of firms' market power in the labor market, as measured by markdowns. We then discuss various robustness checks and heterogeneity patterns.

5.1 Baseline Results

Columns (1) to (3) of Table 1 present the results from OLS, IV, and reduced-form IV estimations, respectively. Across all columns, we observe that migration inflows significantly increase firms' labor market power.

Column (1) reports the OLS estimates, indicating a positive association between immigration and changes in firms' markdowns. Based on the IV estimates in Column (2), an exogenous increase in immigration driven by historical migration networks leads to a statistically significant increase in firms' markdowns. Quantitatively, a 10-percentage-point increase in the immigration rate raises the change in markdown between 2000 and 2006 by approximately 0.026 units. Consistent with this magnitude, comparing prefectures at the 75th (0.096) and 25th (0.025) percentiles of the actual immigration rate implies a difference in markdown growth of about 0.019 units. Relative to a mean change in markdown of 0.666

over the same period, this corresponds to an increase of about 2.9 percent. The reduced-form results in Column (3) are consistent with the IV estimates, indicating that negative exogenous agricultural income shocks induce rural-to-urban migration, which in turn increases firms' monopsony power in local labor markets.

Table 1 Migrants and Markdown: Baseline Results

Dependent variable: $\Delta markdown$			
	(1) OLS	(2) IV	(3) Reduced-form IV
Migration	0.296***	0.259**	
	(0.074)	(0.131)	
IV			-0.891*
			(0.462)
Observations	29,616	29,616	29,616
F-statistic (first stage)		21.95	
R^2	0.005		0.001

Note: Robust standard errors clustered at the prefecture level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample includes firms observed in all years from 2000 to 2006 in the NBS firm census. *Migration* refers to the cumulative immigration rate from 2000–2005, normalized by the baseline (year 2000) destination population. Δmarkdown denotes the change in firm-level markdown over the same period. Columns (1) to (3) report results from OLS, two-stage least squares (2SLS), and reduced-form IV regressions, respectively.

5.2 Robustness

Identification threats We interpret our estimates as capturing the effect of immigration on firms' labor market power. A potential concern is that the shift-share instrument—constructed from agricultural income shocks and historical migration patterns—may exert independent influences on labor market conditions in urban areas, thereby violating the exclusion restriction. First, if rural residents consume manufactured goods produced in urban areas, agricultural income shocks could indirectly affect urban labor demand through changes in product demand. To mitigate this channel, we exclude all migration flows between prefectures located less than 300 kilometers apart in Column 1 of Table A.2. Under the assumption of a gravity model of trade, this restriction helps limit rural-to-urban spillovers via goods markets, ensuring that identification stems from plausibly exogenous variation in migration patterns. In Column 2 of Table A.2, we restrict the sample to exporting firms at

baseline, which are arguably less sensitive to local demand fluctuations and thus less affected by domestic migration-induced demand shocks.

The shift-share design: the shift We explore several alternative specifications of the shift-share instrument to assess the robustness of our main results. Specifically, we consider: (i) restricting the set of crops included in the instrument; (ii) employing an AR(2) process for commodity prices instead of the baseline AR(1) specification; and (iii) isolating exogenous price innovations using the Hodrick–Prescott (HP) filter. We re-estimate the baseline regressions presented in Table 1 using these alternative shift definitions and report the results in Column 3 to 5 of Table A.2. As shown, our key findings remain qualitatively and quantitatively robust across these alternative specifications, suggesting that the estimated effects are not sensitive to the particular construction of the shift component in the shift-share design.

The shift-share design: the share Columns 1 of Table A.3 demonstrate that our main results are robust to alternative constructions of the "share" component in the shift-share instrument. Specifically, we re-estimate the baseline specification using a reallocation of migrants across prefectures along a gravity model. The estimated effects remain consistent with our main findings, reinforcing the validity of the instrument under different assumptions about historical settlement patterns.

Outliers Columns 2 and 3 of Table A.3 further demonstrate that our main results are robust to censoring. Specifically, we re-estimate the regressions using the actual measure of immigration rates, and the actual firm-level markdown respectively. Across both specifications, the estimated effects remain stable, suggesting that our findings are not driven by extreme values or outliers in the data.

Definition of migrant flows Table A.4 presents robustness checks using alternative definitions of migration spells. Specifically, we consider: (i) all migration spells regardless of motive (column 1); (ii) male migrants only (column 2); (iii) individuals with secondary education or less (column 3); and (iv) migration flows not adjusted for return migration (column 4). Across all these alternative definitions, our main results remain robust, indicating that the estimated effects are not sensitive to the particular sample used to define migrant flows.

Randomly generated migration supply shocks To further ensure that our empirical findings are not driven by unobserved confounding factors, we perform a placebo exercise in which we randomly assign migration supply shocks across cities. Specifically, we draw

500 sets of placebo instrumental variables from the empirical distribution of the original shock measure and re-estimate the baseline IV specification in Column (2) of Table 1 for each draw. Figure A.4 presents the distribution of the 500 placebo estimates. The placebo coefficients are centered around a mean of –0.587 with a large dispersion (standard deviation = 11.396), and most of the simulated estimates are statistically insignificant. Therefore, randomly constructed migration shocks do not generate systematic effects on markdowns, suggesting that our main results are unlikely to be driven by unobserved factors or omitted shocks.

5.3 Heterogeneity Analysis

This section explores the heterogeneous effects of migrants on firm's markdowns. Specifically, we examine how the impact of migrant inflows varies by ownership type, export status, capital intensity, and market concentration. These dimensions are critical as they determine the diverse strategic responses of firms to labor supply shocks.

SOEs versus Non-SOEs State-owned enterprises (SOEs) may be subject to non-market constraints (Hsieh and Song, 2015; Cao et al., 2023), causing them to be less responsive to fluctuations in labor market tightness, particularly within the low-skill segment. Columns (1)–(2) of Table 2 split the sample into state-owned enterprises (SOEs) and non–SOEs. We define SOEs as firms in which the ownership share of the state exceeds 25 percent (Poncet et al., 2010). [8] We find that the positive effect of migration on the change in markdown is concentrated among non–SOEs, whereas the coefficient for SOEs is smaller and imprecisely estimated. This pattern is consistent with the view that SOEs face weaker profit-maximization incentives and greater constraints in wage and employment setting (e.g., employment stabilization mandates and preferential access to factor inputs), which dampen their monopsony response to labor-supply shocks.

Exporting firm heterogeneity Exporting firms benefits strongly from the increasing rural-to-urban migration during our sample period (Facchini et al., 2019; Ma and Tang, 2020; Liu and Ma, 2023). Columns (3)–(4) of Table 2 show that the positive impact of migration on markdowns is primarily concentrated among exporting firms. This pattern is consistent with the view that exporting firms are more sensitive to labor supply shocks due to their stronger reliance on labor-intensive production and their higher exposure to international competitive pressure. Access to an expanded pool of low-skilled migrant workers allows exporting firms

[8] The results are robust to defining state ownership using a 50 percent threshold.

to exercise greater monopsony power. In contrast, non-exporting firms typically operating in more localized and protected product markets exhibit relatively limited adjustments in wage-setting power in response to migration inflows.

Capital intensity heterogeneity Columns (5)–(6) of Table 2 investigate heterogeneity by capital intensity, defined as the ratio of real capital stock to employment; firms above the median are classified as "high capital intensity" and those below as "low capital intensity." The estimated effect is larger and statistically significant only for low–capital–intensity firms. This is in line with migrants predominantly entering labor-intensive production where switching costs and contracting frictions are more salient, yielding a lower firm-level labor-supply elasticity and hence a larger increase in markdowns when labor supply expands.

Market concentration heterogeneity Columns (7)–(8) of Table 2 show that the effect of migration on markdowns is more pronounced in high-concentration industries. In sectors with weaker product-market competition, firms have greater ability to translate labor supply increases into wage suppression, resulting in larger markdown responses. This suggests that greater product-market power reinforces firms' ability to suppress wages when labor supply expands.

Table 2 Heterogeneity Analysis

Dependent variable: $\Delta markdown$				
	(1) SOEs	(2) Non-SOEs	(3) Exporting firms	(4) Non-exporting firms
migration	0.255	0.309**	0.269*	0.264
	(0.332)	(0.136)	(0.150)	(0.253)
Observations	4,423	25,193	11,249	18,367
	(5) High capital intensity	(6) Low capital intensity	(7) High concentration	(8) Low concentration
migration	0.226	0.305*	0.281**	0.243
	(0.140)	(0.158)	(0.116)	(0.166)
Observations	12,798	16,812	12,761	16,855

Note: Robust standard errors clustered at the prefecture level are in parentheses. Columns (1)–(2) divide the sample by ownership (SOEs vs. non-SOEs), based on whether firms' share of state-owned capital is more than 25 percent. Columns (3)–(4) split the sample by exporting status, based on whether firms report positive export sales in the pre-period. Columns (5)–(6) divide by capital intensity (high vs. low), defined as real fixed capital over employment; firms above (below) the median capital intensity in 2000 are classified as high (low) capital intensity. Columns (7)–(8) split by industry-level market concentration (high vs. low), measured using the Herfindahl–Hirschman Index (HHI) computed from the market shares of the top 50 firms in industrial sales within each 2-digit industry; industries above (below) the median HHI are classified as high (low) concentration. All classifications are based on firms' pre-period (year 2000) characteristics. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6 Mechanisms

We next formally investigate the underlying mechanisms through which exogenous migrant inflows facilitate an expansion in firm-level markdowns. Specifically, we elucidate the channels that translate these labor supply shocks into increased monopsony power. Guided by our theoretical framework, we proceed in three stages. First, we establish that firms with a higher intensity of low-skilled labor exhibit structurally higher wage markdowns. More importantly, we demonstrate that the markdowns of these firms are disproportionately amplified by increasing migrant inflows, as Proposition 2 predicts. Second, we provide evidence

that migrant inflows lead to a significant decline in skill intensity—measured by various metrics—within both local labor markets and individual firms. This suggests a systematic shift toward low-skill intensive production following the labor supply shock. Finally, and most crucially, we demonstrate that migrant inflows lead to a significant reduction in the residual labor supply elasticities facing firms. Given that low-skilled workers generally exhibit lower labor supply elasticities than their high-skilled counterparts, an exogenous influx of rural-to-urban migrants shifts the aggregate labor composition toward the less elastic segment. This downward pressure on effective labor supply elasticity drives the increase in firm markdowns, providing empirical evidence that aligns precisely with our theoretical predictions.

6.1 Migrants, Skill Intensities and Markdown

Skill Intensities and Firms' Markdown We now examine how firms' wage-setting power in the labor market varies with their skill intensity. Low-skilled workers are closer substitutes for migrants than high-skilled workers, firms with a lower share of high-skilled employees may rely more heavily on migrant labor and face a more inelastic labor supply. To the extent that migrant inflows amplify labor market power, our theory hypothesizes that firms with a greater reliance on low-skilled labor should exhibit higher wage markdowns.

To test this hypothesis, we draw on data from the 2004 Economic Census and estimate the following cross-sectional regression:

$$\text{markdown}_i = \alpha + \beta \cdot \text{employment_share}_i + \lambda_{\text{province}} + \lambda_{\text{industry}} + \varepsilon_i \tag{28}$$

In this specification, the dependent variable is the wage markdown, which we estimate using the previously described procedure[9], and the key explanatory variable is *employment_share*, which captures a firm's employment structure by education level. To ensure the robustness of our results, we define skill intensity using the employment shares of four distinct cohorts: (i) those with at most a high school diploma, (ii) associate degree holders, (iii) bachelor's degree holders, and (iv) those with post-graduate qualifications. All regressions include province fixed effects and two-digit industry fixed effects. Standard errors are clustered at the firm level.

Estimation results using the above specification are reported in Table 3. It indicates that firms with a higher concentration of low-skilled workers—specifically those with a high school education or less—exhibit significantly higher markdowns. This skill-markdown gradient reverses for more educated cohorts, where the share of employees with post-secondary degrees

[9] We rely on the Firm Census Data to run this regression since the ASIF does not provide detailed information on firm's skill composition.

is associated with compressed markdowns. These results provides direct empirical validation for the foundational assumption of our theoretical model and establishes the basis for the heterogeneous impact of migrant inflows across firms with varying skill intensities.

Table 3 Skill Intensity and Markdown

Dependent variable: *markdown*				
	(1)	(2)	(3)	(4)
senior_below	0.695***			
	(0.082)			
associate		-0.698***		
		(0.121)		
bachelor			-1.872***	
			(0.197)	
graduate				-1.960**
				(0.817)
Province FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	24,084	24,084	24,084	24,084
Adjusted R^2	0.404	0.402	0.405	0.401

Note: Robust standard errors clustered at the prefecture level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The regression uses 2004 cross-sectional data, based on a balanced panel of firms observed continuously from 2000 to 2006 in the NBS firm census. *markdown* is the firms' labor market power measure. Columns (1) through (4) report results using the share of employees with (i) high school education or below, (ii) associate degrees, (iii) bachelor's degrees, and (iv) graduate education or above as the independent variable, respectively.

Heterogeneous Effect of Skill Intensity on Markdown When Migrants Increases While firms with a larger share of low-skilled workers exhibit structurally higher baseline markdowns, Proposition 2 further establishes that these firms experience a disproportionately larger increase in markdowns following an influx of migrants. This subsection investigates whether the impact of migration inflows on firms' markdowns varies with the skill intensity of the local labor force. Consistent with the implications derived from our theoretical framework, we expect that firms operating in more low-skill–intensive environments benefit more from migration-induced labor supply expansion, thereby exhibiting stronger monopsony power.

To empirically examine this mechanism, we construct measures of skill intensity at three

different levels, inclusing the firm level, the city level, and the industry level. In all cases, skill intensity is defined as the share of employees (or migrants) holding an associate degree or higher. We incorporate these measures into our baseline regression specification by interacting migration inflows with the corresponding skill intensity variable.

Table 4 presents the results. Across all three levels of measurement, the interaction terms are negative and statistically significant, indicating that the effect of migrant inflows are more pronounced in low-skill–intensive settings. These results align closely with Proposition 2 of our theoretical framework that the positive effect of migrants on firms' markdown would be stronger for firms with higher low-skill technologies.

Table 4 The Heterogeneous Effect of Skill Intensity

Dependent variable:	*markdown*		
	(1) Firm level	(2) City level	(3) Industry level
Migration	0.844***	-0.071	0.570***
	(0.214)	(0.166)	(0.172)
Migration × skill intensity	-0.371***	-4.773*	-4.708*
	(0.109)	(2.619)	(2.546)
Observations	11,140	27,157	28,102

Note: Robust standard errors clustered at the prefecture level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Columns (1)–(3) present results based on firm-level, city-level, and industry-level measures of skill intensity, respectively. Skill intensity is defined as the share of migrants (or employees) with an associate degree or higher.

6.2 Migrants and Firms' Labor Composition

The heterogeneous impact on firm-level markdowns is driven by internal labor reallocation, specifically a systematic change in the skill mix of the workforce. Consequently, we formally ascertain whether migrant inflows drive these observed shifts in firm-level labor composition. Specifically, we investigate whether destinations exposed to larger migrant shocks exhibit a disproportionate expansion of low-skilled employment relative to high-skilled labor, thereby precipitating a fundamental reconfiguration of the firm's employment structure.

We begin by considering whether migration changes the skill composition of the migrant population itself at the destination cities. Using data from the 2005 1% Population Survey, we construct a city-level measure of labor composition among migrants, defined as the share of migrant workers holding an associate degree or above. We compute the change in this measure between 2000 and 2005 and regress it on cumulative migration inflows to the city. The results, reported in Column (1) of Table 5, indicate that destinations experiencing

larger migrant inflows receive disproportionately more low-skilled migrants. This provides direct evidence that internal migration in China contributed to a compositional shift toward lower-educated labor in migrant-receiving cities.

We then turn to firm-level evidence. Using matched data from the 1995 and 2004 Economic Censuses, we compute changes in firms' employment structure over this period, including the change in the share of high-skilled employees, and the change in the log number of high-skilled and low-skilled workers, and replace the dependent variable in our baseline specification (23) accordingly. The results are reported in Columns (2)–(4) of Table 5. Consistent with the findings based on migrant population data, we find that destinations receiving larger migrant inflows experience a statistically significant decline in the share of high-skilled workers employed by firms. As shown in Column (2) of Table 5, a 10-percentage-point increase in the immigration rate is associated with a 0.55-percentage-point reduction in the high-skill employment share.

However, when turning to changes in the number of workers, we find that migration inflows lead to increases in both high-skilled and low-skilled employment. As shown in Columns (3) and (4) of Table 5, a 10-percentage-point rise in the immigration rate is associated with a 4.6 percent increase in high-skilled employment and a substantially larger 9.1 percent increase in low-skilled employment.[10]

We next examine how migration inflows affect wages at the firm and prefecture levels, with the aim of shedding light on the mechanisms through which migration-induced changes in labor composition translate into adjustments in labor market power. The corresponding results are reported in Table A.5. As shown in Column (1), at the firm level, we find that immigration significantly reduces average wages, consistent with a compositional shift toward lower-paid workers. At the prefecture level, using wage measures constructed from the Urban Household Survey (UHS),[11] we find that immigration is associated with a decline in residents' average wages, although the effect is not statistically significant (see column 2). This muted wage response among residents is consistent with several mechanisms emphasized in Imbert et al. (2022), including partial segmentation of labor markets, greater wage rigidity for incumbent resident workers, and the possibility that the recruitment of lower-productivity migrants lowers firms' average labor costs without proportionately affecting residents' wages.

[10] These empirical results suggest that workers of different skill levels are imperfect substitutes, corroborating the core assumption of our theoretical model in equation (4). Consequently, an influx of low-skilled migrants induces firms to hire a disproportionately larger number of low-skilled workers relative to high-skilled workers.

[11] A detailed description of the Urban Household Survey (UHS) is provided in Appendix B.1. It is worth noting that the UHS covers only a subset of provinces in China. Specifically, the survey includes 16 provinces and municipalities, including Beijing, Shanxi, Liaoning, Heilongjiang, Shanghai , Jiangsu, Anhui, Jiangxi, Shandong, Henan, Hubei, Guangdong, Chongqing, Sichuan, Yunnan, and Gansu.

Importantly, decomposing prefecture-level resident wages by skill group reveals pronounced heterogeneity. Immigration significantly reduces the average wages of low-skilled resident workers, whereas the estimated effect on high-skilled residents' wages is statistically insignificant and positive in sign (see columns (3) and (4) of Table A.5). These results point to an increase in competitive pressure among workers at the lower end of the skill distribution. Migration inflows intensify competition primarily among low-skilled workers, where wage adjustments are more pronounced, while high-skilled workers remain largely insulated from direct wage effects.

Taken together, these findings show that migration-driven labor supply shocks reshape the internal labor allocation of firms by disproportionately expanding the employment of low-skilled labor, thereby shifting firms' labor composition toward low-skilled workers. Given that low-skilled workers face systematically lower labor supply elasticities than high-skilled workers—a premise we formally test in the subsequent section—the aforementioned shifts in labor composition necessarily alter a firm's aggregate labor supply elasticity. This, in turn, induces a predictable change in markdowns, as the firm-level markdown is characterized by a one-to-one mapping with its aggregate labor supply elasticity.

Table 5 Migrants and Labor Composition

Dependent variable:	Share of high skilled migrants	Share of high skilled employees	Number of high skilled employees	Number of low skilled employees
	(1)	(2)	(3)	(4)
migration	-0.122***	-0.055***	0.460**	0.907***
	(0.044)	(0.014)	(0.197)	(0.228)
Observations	245	36,061	36,061	36,061

Note: Robust standard errors clustered at the prefecture level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Column (1) presents results based on city-level measures of labor composition, while Columns (2) to (4) use firm-level measures to investigate the effect of migration on firms' labor composition. High-skilled migrants (or employees) are defined as those with an associate degree or higher.

6.3 Migrants and Labor Supply Elasticities

Evidence of Low Labor Supply Elasticities For Low-Skilled Workers Having established at the firm level that migration inflows make labor supply more inelastic, we now examine whether this effect varies across different types of workers. In particular, because migration inflows disproportionately expand the pool of low-skilled labor, it is plausible that

firms gain greater monopsony power over low-skilled workers than over high-skilled ones. To investigate this heterogeneity, we estimate labor supply elasticities separately for high- and low-skilled workers using bilateral migration flows and the corresponding wage responses.

This exercise complements the firm-level evidence by moving from the elasticity faced by a single firm to supply responsiveness at the city–skill group level, allowing us to explicitly trace how migration affects wage-setting power across different segments of workers. By doing so, we document whether migration-driven labor supply shocks strengthens firms' monopsony power over low-skilled workers , the group most directly affected by rural-to-urban migration flows.

Following Khanna et al. (2025), we estimate skill–specific labor supply elasticities from a bilateral choice equation that maps the share of migrants moving from origin city o to destination city d onto destination wages and migration costs. For each skill group $s \in \{\text{low}, \text{high}\}$, let π_{sod} denote the share of movers from (s, o) who choose destination d. We estimate:

$$\log \pi_{sod} = \eta_s \log w_{sd} + X'_{od}\beta_x + \varepsilon_{sod}. \tag{29}$$

where $\log w_{sd}$ is the wage at destination d for skill group s; X_{od} includes additional bilateral frictions as needed The vector X_{od} collects additional bilateral frictions, including the great–circle distance between origin o and destination d, origin–city fixed effects, and the hukou restriction index of destination d from Fan (2019). The coefficient on $\log w_{sd}$ is η_s, which is the labor supply elasticity for skill group s.

To trace out the labor supply curves we need exogenous shifts in labor demand. We instrument $\log w_{sd}$ with a Bartik-style shock built from changes in *world import demand* (WID) at the industry level (Autor et al., 2013). Intuitively, exogenous variation in global demand for the products a city is specialized in raises local labor demand and therefore equilibrium wages, allowing us to identify the supply elasticity.

For each skill group $s \in \{\text{low}, \text{high}\}$ and city d, we construct

$$\text{WID}_d^s \;=\; \sum_i \omega_{sdi}^{1998} \cdot \frac{\text{World IM}_{i,2006} - \text{World IM}_{i,2000}}{\text{World IM}_{i,2000}}, \qquad \omega_{sdi}^{1998} \equiv \frac{EX_{sdi}^{1998}}{\sum_j EX_{sdj}^{1998}}, \tag{30}$$

where i indexes CIC 4-digit industries; EX_{sdi}^{1998} is city d's base-year (1998) exports in industry i, so that ω_{sdi}^{1998} are pre-determined export shares; and $\text{World IM}_{i,t}$ denotes world imports of industry i at year t.

When instrumenting the *high-skill* wage we restrict i to skill-intensive industries and get $\text{WID}_d^{\text{high}}$. Similarly, we build $\text{WID}_d^{\text{low}}$ using non–skill-intensive industries to instrument the *low-skill* wage.[12] For wages of high- and low-skill workers, we use microdata from the Urban

[12] Using the 1995 Industrial Census, we first aggregate firm-level employment by education to the four-

Household Survey (UHS), with details provided in Appendix B.1; the world import-demand shifters are constructed from the UN Comtrade database.

Table 6 Labor supply elasticities by skill group

Dependent variable:	**Employment level**		**Employment share**	
	(1) Low skill	(2) High skill	(3) Low skill	(4) High skill
Log(wage)	0.423***	0.862***	1.422***	2.653***
	(0.036)	(0.073)	(0.193)	(0.262)
Log(distance)	-0.155***	-0.126***	-0.644***	-0.383***
	(0.008)	(0.006)	(0.033)	(0.021)
hukou	-0.001	-0.080***	0.106***	-0.211***
	(0.004)	(0.010)	(0.025)	(0.036)
Origin-city FE	Yes	Yes	Yes	Yes
Observations	47,460	47,460	47,460	47,460
F statistics	193657.163	106011.889	193657.163	106011.889

Note: Robust standard errors clustered at the prefecture level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6 reports 2SLS estimates of equation (29) using the Bartik instrument in equation (30). Across specifications that alternately use the *employment level* (cols. 1–2) and the *employment share* (cols. 3–4) as the dependent variable, the estimated wage elasticities of migration are consistently larger for high-skill workers than for low-skill workers, indicating that high-skill migrants respond more strongly to wage differentials, whereas low-skill movers exhibit a more inelastic response.

Effect of Increasing Migrants on Firm Level Labor Supply Elasticities We finally turn to the mechanism linking migration-induced labor composition changes to firms' monopsony power. In our theoretical framework, Equation (3) establishes a direct one-to-one relationship between the labor supply elasticity of migrant workers and firms' markdowns: a lower labor supply elasticity implies a steeper supply curve facing the firm, increasing its monopsony power and thus raising markdowns. This motivates an empirical test of whether migration inflows affect the labor supply elasticities that firms face.

To do so, we follow the methodology proposed by Rubens et al. (2024) and estimate firm-level (inverse) labor supply elasticities, which measure the responsiveness of wages to changes in firm employment. Appendix B.3 provides detailed procedures about firm level labor

digit CIC industry level. We then compute each industry's skill intensity, defined as the share of skilled workers (associate degree or above) in total employment, and classify industries as "skill-intensive" if their skill intensity is above the cross-industry median.

supply elasticity estimation. Intuitively, a larger inverse elasticity corresponds to a smaller direct elasticity, indicating that labor is less responsive to wage changes, i.e., supply is more inelastic. A more inelastic labor supply curve enables firms to exercise greater monopsony power, suppress wages further below marginal revenue product, and thus generate higher markdowns.

Table 7 provides strong empirical support for this mechanism. Both OLS and IV estimates show that destinations experiencing higher migration inflows exhibit significantly greater inverse labor supply elasticities, meaning that migration makes labor supply to firms more inelastic. This increased inelasticity consistent with the rise in markdowns documented earlier, reflecting greater monopsony power at the firm level.

Table 7 Migrants and Labor Supply Elasticity

Dependent variable:	Δ*inverse labor supply elasticity*		
	(1)	(2)	(3)
	OLS	IV	Reduced-form IV
Migration	9.195***	11.838**	
	(3.329)	(5.501)	
IV			-39.982**
			(16.030)
Observations	50,004	50,004	50,004
F-statistic (first stage)		23.38	
R^2	0.0004		0.0002

Note: Robust standard errors clustered at the prefecture level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Columns (1)–(3) report OLS, 2SLS, and reduced-form IV estimates, respectively.

Taken together with the evidence on firm-level inverse elasticities, migration inflows that are disproportionately *low-skill* make the labor supply facing firms more inelastic. A more inelastic supply curve expands a firm's scope to exercise monopsony power and compress wages, whereas a higher elasticity (as for high-skill workers) limits such wage-setting power.

7 Conclusion

This paper offers the first integrated analysis, both theoretically and empirically, of how rising migrant inflows shape firms' monopsony power. We first build a theoretical model, in which migration is an endogenous choice driven by the gap between agricultural returns

and industrial wages. Manufacturing firms utilize a combination of tasks with varying skill intensities. Since low-skilled labor is characterized by a more inelastic supply, an exogenous migration shock alters the internal labor-mix and depresses the aggregate elasticity facing the firm. Consequently, firms exercise greater wage-setting power, leading to higher markdowns, especially within sectors predominantly employing low-skill intensive technologies.

We test our theoretical propositions by exploiting exogenous variation in agricultural income shocks within a Bartik-style instrumental variable framework. Our empirical results indicate that firms with a higher share of low-skilled workers exercise greater monopsony power, as evidenced by higher wage markdowns. An influx of low-skilled migrants reinforces this pattern by lowering the aggregate labor supply elasticity facing the firm. Heterogeneity analysis reveals that the capacity for wage suppression is more pronounced among non-SOEs, exporters, and firms in industries characterized by lower capital intensity and higher product market concentration. Overall, our empirical evidence provides strong support for the model's predictions. Our study underscores the necessity of understanding the structural impact of migration on firm behavior. Specifically, we emphasize that the compositional heterogeneity of the labor pool plays a critical role in determining the welfare implications of exogenous labor market shocks. Neglecting these mechanisms would lead to a biased and incomplete assessment of the welfare implications of migration.

References

Accetturo, Antonio, Matteo Bugamelli, and Andrea Roberto Lamorgese, "Welcome to the machine: firms' reaction to low-skilled immigration," *Bank of Italy Temi di Discussione (Working Paper) No*, 2012, *846.*

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 2011, *4b* (16082), 1043–1171.

Alan, Manning, "Imperfect competition in the labor market," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 973–1041.

Amior and Manning, "Monopsony and the Wage Effects of Migration," *The Economic Journal*, 2025.

Angelucci, Manuela, "Migration and Financial Constraints: Evidence from Mexico," *Review of Economics Statistics*, 2015, *97* (1), 224–228.

Ashenfelter, Orley C, Henry Farber, and Michael R Ransom, "Labor market monopsony," *Journal of Labor Economics*, 2010, *28* (2), 203–210.

Autor, David H. and Levy Richard J. Murnane, "The skill content of recent technological change: an empirical exploration," *The Quarterly Journal of Economics*, 2003, *118* (4), 1279–1333.

Autor, David H, David Dorn, and Gordon H Hanson, "The China syndrome: Local labor market effects of import competition in the United States," *American economic review*, 2013, *103* (6), 2121–2168.

Behrens, Kristian and Yasuhiro Sato, "Migration, skill formation, and the wage structure," *Journal of Regional Science*, 2011, *51* (1), 5–30.

Berger, David, Kyle Herkenhoff, and Simon Mongey, "Labor market power," *American Economic Review*, 2022, *112* (4), 1147–1193.

Berry, Steven T, "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 1994, pp. 242–262.

Borjas, George J, "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market," *The quarterly journal of economics*, 2003, *118* (4), 1335–1374.

Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang, "WTO accession and performance of Chinese manufacturing firms," *American Economic Review*, 2017, *107* (9), 2784–2820.

Brücker, Herbert and Elke J Jahn, "Migration and wage-setting: reassessing the labor market effects of migration," *Scandinavian journal of economics*, 2011, *113* (2), 286–317.

Bryan, Gharad and Melanie Morten, "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

Bustos, Paula, Juan Manuel Castro Vincenzi, Joan Monras, and Jacopo Ponticelli, "Structural Transformation, Industrial Specialization, and Endogenous Growth," *NBER Working Papers*, 2019.

Cao, Yiming, Raymond Fisman, Hui Lin, and Yongxiang Wang, "SOEs and Soft Incentive Constraints in State Bank Lending," *American Economic Journal: Economic Policy*, 2023, *15* (1), 174–195.

Card, David, "The impact of the Mariel boatlift on the Miami labor market," *Ilr Review*, 1990, *43* (2), 245–257.

_ , Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 2018, *36.*

Dustmann, Christian, Tommaso Frattini, and Ian P Preston, "The effect of immigration along the distribution of wages," *Review of Economic Studies*, 2013, *80* (1), 145–173.

Facchini, Giovanni, Maggie Y. Liu, Anna Maria Mayda, and Minghai Zhou, "China's "Great Migration": The Impact of the Reduction in Trade Policy Uncertainty," *Journal of International Economics*, 2019, *120*, 126–144.

Fan, Jingting, "Internal geography, labor mobility, and the distributional impacts of trade," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 252–288.

Feenstra, Robert C. and Chang Hong, "China's Exports and Employment," NBER Working Paper 13552 October 2007.

Feng, Shuaizhang, Yingyao Hu, and Robert Moffitt, "Long Run Trends in Unemployment and Labor Force Participation in China," *Journal of Comparative Economics*, 2017, *45* (2), 304–324.

Foged, Mette and Giovanni Peri, "Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data," *American Economic Journal: Applied Economics*, 2016, *8.*

Gai, Qingen, Naijia Guo, Bingjing Li, Qinghua Shi, and Xiaodong Zhu, "Migration Costs, Sorting, and the Agricultural Productivity Gap," *Working Paper*, 2021.

Garriga, Carlos, Aaron Hedlund, Yang Tang, and Ping Wang, "Rural-Urban Migration, Structural Transformation, and Housing Markets in China," *American Economic Journal: Macroeconomics*, 2023, *15* (2), 413–440.

Ge, Suqin and Dennis Tao Yang, "Changes in China's Wage Structure," *Journal of the European Economic Association*, 2014, *12* (2), 300–336.

Gollin, Douglas, Stephen Parente, and Richard Rogerson, "The Role of Agriculture in Development," *American Economic Review*, 2002, *92* (2), 160–164.

Gray, Rowena, Giulia Montresor, and Greg C Wright, "Processing immigration shocks: Firm responses on the innovation margin," *Journal of International Economics*, 2020, *126*, 103345.

Groeger, Andre and Yanos Zylberberg, "Internal Labor Migration as a Shock Coping Strategy: Evidence from a Typhoon," *American Economic Journal: Applied Economics*, 2016, *8* (2), 123–153.

Grossman, Gene M. and Esteban Rossi-Hansberg, "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, 2008, *98.*

Grossmann, V. and D. Stadelmann, "Wage Effects of High-skilled Migration: International Evidence," *Social Science Electronic Publishing*, 2013, *27* (2), 297–319.

Herrendorf, Berthold, Christopher Herrington, and Akos Valentinyi, "Sectoral Technology and Structural Transformation," *American Economic Journal: Macroeconomics*, 2013, *7* (4), 104–133.

Hsieh, Chang-Tai and Zheng Song, "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China," *Brookings Papers on Economic Activity*, 2015, *2015* (1), 295–366.

Huang, Jikun, Will Martin, Yu Liu, and Scott Rozelle, "Agricultural Trade Reform and Rural Prosperity: Lessons from China," *Food Politics & Sociology eJournal*, 2008.

Imbert, Clément and John Papp, "Short-term migration, rural public works, and urban labor markets: Evidence from India," *Journal of the European Economic Association*, 2020, *18* (2), 927–963.

_ , Marlon Seror, Yifan Zhang, and Yanos Zylberberg, "Migrants and Firms: Evidence from China," *American Economic Review*, 2022, *112* (6), 1885–1914.

Kaivan, Munshi, "Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market," *Quarterly Journal of Economics*, 2003, *118* (2), págs. 549–599.

Kerr, Sari Pekkala, William R Kerr, and William F Lincoln, "Skilled immigration and the employment structures of US firms," *Journal of Labor Economics*, 2015, *33* (S1), S147–S186.

Khanna, Gaurav, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song, "The productivity consequences of pollution-induced migration in China," *American Economic Journal: Applied Economics*, 2025, *17* (2), 184–224.

Kleemans, Marieke and Jeremy Magruder, "Labour market responses to immigration: Evidence from internal migration driven by weather shocks," *The Economic Journal*, 2018, *128* (613), 2032–2065.

Kondo, Illenin O., Yao Amber Li, and Wei Qian, "Trade liberalization and labor monopsony: Evidence from Chinese firms," *Journal of International Economics*, 2024, *152*.

Lewis, Ethan, "Immigration, skill mix, and capital skill complementarity," *The Quarterly Journal of Economics*, 2011, *126* (2), 1029–1069.

Lewis, William Arthur, "Economic Development with Unlimited Supplies of Labour," *Manchester School*, 1954, *22* (2), 139–191.

Liu, Chen and Xiao Ma, "Migration, tariffs, and China's export surge," *Journal of International Economics*, 2023, *140*, 103713.

Loecker, Jan De and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, 2012, *102* (6), 2437–2471.

Long, Cheryl and Xiaobo Zhang, "Cluster-based industrialization in China: Financing and performance," *Journal of international economics*, 2011, *84* (1), 112–123.

Lopez, Ramon and Maurice Schiff, "Migration and the skill composition of the labour force: The impact of trade liberalization in LDCs," *Canadian Journal of Economics*, 1998, *31* (2).

Ma, Lin and Yang Tang, "Geography, trade, and internal migration in China," *Journal of Urban Economics*, 2020, *115*, 103181.

Manning, Alan, "The real thin theory: monopsony in modern labour markets," *Labour Economics*, 2003, *10* (2), 105–131.

_ , "The real thin theory: monopsony in modern labour markets," *Labour economics*, 2003, *10* (2), 105–131.

_ , "Imperfect Competition in the Labor Market," *Handbook of Labor Economics*, 2011, *4*, 973–1041.

Minale, Luigi, "Agricultural productivity shocks, labour reallocation and rural-urban migration in China," *Journal of economic geography*, 2018, *18* (4), 795–821.

Monras, Joan, "Immigration and wage dynamics: Evidence from the mexican peso crisis," *Journal of Political Economy*, 2020, *128* (8), 3017–3089.

Morten, Melissa, "Temporary Migration and Endogenous Risk Sharing in Village India," *Journal of Political Economy*, 2019, *127* (1), 1–46.

Mortensen, Dale T. and Kenneth Burdett, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 1998, *39* (2), 257–273.

Ottaviano, Gianmarco I. P., Giovanni Peri, and Greg C. Wright, "Immigration, Offshoring and American Jobs," *American Economic Review*, 2013, *103* (5), 1925–1959.

Ottaviano, Gianmarco IP and Giovanni Peri, "Rethinking the effect of immigration on wages," *Journal of the European economic association*, 2012, *10* (1), 152–197.

Pham, Hoang, "Trade reform, oligopsony, and labor market distortion: Theory and evidence," *Journal of International Economics*, 2023, *144* (000), 39.

Poncet, Sandra, Walter Steingress, and Hylke Vandenbussche, "Financial constraints in China: Firm-level evidence," *China Economic Review*, 2010, *21* (3), 411–422.

Rubens, Michael, "Market Structure, Oligopsony Power, and Productivity," *American Economic Review*, 2023, *113* (9), 2382–2410.

_ , **Yingjie Wu, and Mingzhi (Jimmy) Xu**, "Exploiting or Augmenting Labor?," *American Economic Review: Insights*, 2024. Forthcoming.

Sparber, Chad and Giovanni Peri, "Task Specialization, Immigration, and Wages," *American Economic Journal Applied Economics*, 2009, *1* (3), 135–169.

Tombe, Trevor and Xiaodong Zhu, "Trade, migration, and productivity: A quantitative analysis of China," *American Economic Review*, 2019, *109* (5), 1843–1872.

Xie, Enze, Mingzhi Xu, and Miaojie Yu, "Trade liberalization, labor market power, and misallocation across firms: Evidence from China's WTO accession," *Journal of Development Economics*, 2024, *171*, 103353.

Yeh, C., C. Macaluso, and B. Hershbein, "Monopsony in the US Labor Market," *American Economic Review*, 2022, *112* (7), 2099–2138.

Online Appendix

A Additional Figures and Tables

Table A.1 First-Stage Regression: Immigration

Dependent variable:	**Immigration**			
	(1) Inter-prefecture	(2) Outside 100 km	(3) Outside 300 km	(4) Outside 500 km
Shift-share instrument	-1.025***	-0.991***	-0.781***	-0.607***
	(0.173)	(0.169)	(0.150)	(0.119)
Observations	315	315	315	315
Adjusted R^2	0.098	0.096	0.077	0.074

Note: Robust standard errors clustered at the prefecture level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the number of rural migrants originating from (i) other prefectures, or from prefectures located more than (ii) 100 km, (iii) 300 km, and (iv) 500 km away from the destination, respectively, normalized by the number of urban residents in the destination prefecture.

Table A.2 Sensitivity Analysis: Identification and Shift Design

Dependent variable: Δmarkdown					
	(1)	(2)	(3)	(4)	(5)
Migration	0.308**	0.269*	0.424***	0.258*	0.351***
	(0.156)	(0.150)	(0.144)	(0.133)	(0.136)
Observations	29,616	11,249	29,616	29,616	29,616

Note: Robust standard errors clustered at the prefecture level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample includes firms observed in all years from 2000 to 2006 in the NBS firm census. *Migration* refers to the cumulative immigration rate from 2000–2005, normalized by the baseline (year 2000) destination population. Δmarkdown denotes the change in firm-level markdown over the same period. Columns (1) and (2) report results excluding migration within a 300km radius and restricting the sample to exporters, respectively. Column (3) uses a restricted crop list to construct price shocks. Column (4) uses an AR(2) specification for the price shock innovation, while Column (5) applies a Hodrick–Prescott filter (parameter = 14,400) to logged commodity prices.

Table A.3 Sensitivity Analysis: Share Design and Outliers

Dependent variable: Δmarkdown			
	(1)	(2)	(3)
Migration	0.469***	0.150*	0.470*
	(0.157)	(0.089)	(0.265)
Observations	29,616	29,616	29,616

Note: Robust standard errors clustered at the prefecture level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample includes firms observed in all years from 2000 to 2006 in the NBS firm census. *Migration* refers to the cumulative immigration rate from 2000–2005, normalized by the baseline (year 2000) destination population. Δmarkdown denotes the change in firm-level markdown over the same period. Column (1) presents results using an alternative construction of the "share" component in the shift-share instrument based on a gravity-model-based reallocation of migrants. Columns (2) and (3) address outlier sensitivity by winsorizing the immigration rate and markdown measures, respectively.

Table A.4 Sensitivity Analysis: Migrant Flows Definition

Dependent variable: Δmarkdown				
	(1)	(2)	(3)	(4)
	All migrants	Male only	Low-education only	No return adjustment
Migration	0.263**	0.480*	0.281*	0.302**
	(0.132)	(0.254)	(0.149)	(0.151)
Observations	29,616	29,616	29,616	29,616

Note: Robust standard errors clustered at the prefecture level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample includes firms observed in all years from 2000 to 2006 in the NBS firm census. *Migration* refers to the cumulative immigration rate over 2000–2005, normalized by the baseline (year 2000) population in the destination prefecture. Δmarkdown denotes the change in firm-level markdown during the same period. Column (1) includes all migration spells. Column (2) restricts the sample to male migrants; Column (3) to low-education individuals; and Column (4) does not adjust for return migration.

To examine whether the migration-induced changes in labor composition documented in Section 6.2 are also reflected in wage outcomes, we investigate the effects of migration on average wages at both the firm and resident levels. We begin with firm-level regressions using manufacturing firm data, where average wages are constructed as total wage payments divided by total employment at the firm level. This specification replaces the dependent variable in the baseline regression in equation (23), allowing us to assess whether the compositional shift toward low-skilled labor is associated with changes in firms' average wages. As shown in Column (1) of Table A.5, immigration inflows are associated with a decline in firms' average wages, consistent with an increased reliance on lower-paid labor.

We then turn to residents' labor market outcomes using data from the Urban Household Survey (UHS) between 2002 and 2006. To assess whether the wage responses of residents are consistent with the labor composition patterns documented above, we estimate the following specification:

$$\Delta y_d = \alpha + \beta m_d + X_d\gamma + u_d, \tag{A.1}$$

where Δy_d represents the change in average labor market outcomes among surveyed residents in prefecture d over the period 2002-2006. The variable m_d measures the immigration rate between 2001 and 2005 and is instrumented using a shift-share approach based on z_d. The vector X_d captures average demographic characteristics of residents, including age, gender, and marital status, while u_d denotes the error term.

Column (2) of Table A.5 reports the effect of migration on the average wage of all urban residents. While the estimated coefficient is negative, it is not statistically significant, suggesting that the wage adjustment among residents is relatively less elastic. Columns (3) and (4) further decompose resident wages by skill group. The results indicate that immigration significantly reduces the average wages of low-skilled urban residents, whereas the effect on high-skilled residents' wages is statistically insignificant and positive in sign. This pattern is consistent with imperfect substitutability between high- and low-skilled labor and suggests that migration inflows disproportionately affect segments of the labor market characterized by more inelastic labor supply. In particular, competitive pressure is concentrated at the lower end of the skill distribution, where wages adjust more strongly, while high-skilled workers appear relatively insulated from direct wage effects.

Table A.5 Migration and Average Wages

Dependent variable:	**Average wage**			
	(1) Manufacturing firms	(2) All urban residents	(3) Low-skill urban residents	(4) High-skill urban residents
Migration	-0.103*	-0.265	-0.467**	0.140
	(0.062)	(0.165)	(0.210)	(0.182)
Observations	29,616	172	172	172

Note: Robust standard errors clustered at the prefecture level are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the logarithm of average wages. Column (1) reports firm-level regressions using manufacturing firm data, where average wages are constructed as total wage payments divided by total employment at the firm level; this specification replaces the dependent variable in the baseline regression (23). Columns (2)-(4) use prefecture-level wage measures constructed from the Urban Household Survey (UHS) by aggregating individual wages to the prefecture level. In Columns (2)-(4), the dependent variables are constructed as changes in outcomes between 2002 and 2006, and the estimation follows regression equation(A.1). Column (2) uses the average wage of all urban residents. Column (3) restricts the sample to low-skilled urban residents, defined as individuals with educational attainment of high school or below. Column (4) restricts the sample to high-skilled urban residents, defined as those with education above high school. Regressions in Columns (2)-(4) include the same set of controls, namely average individual characteristics.

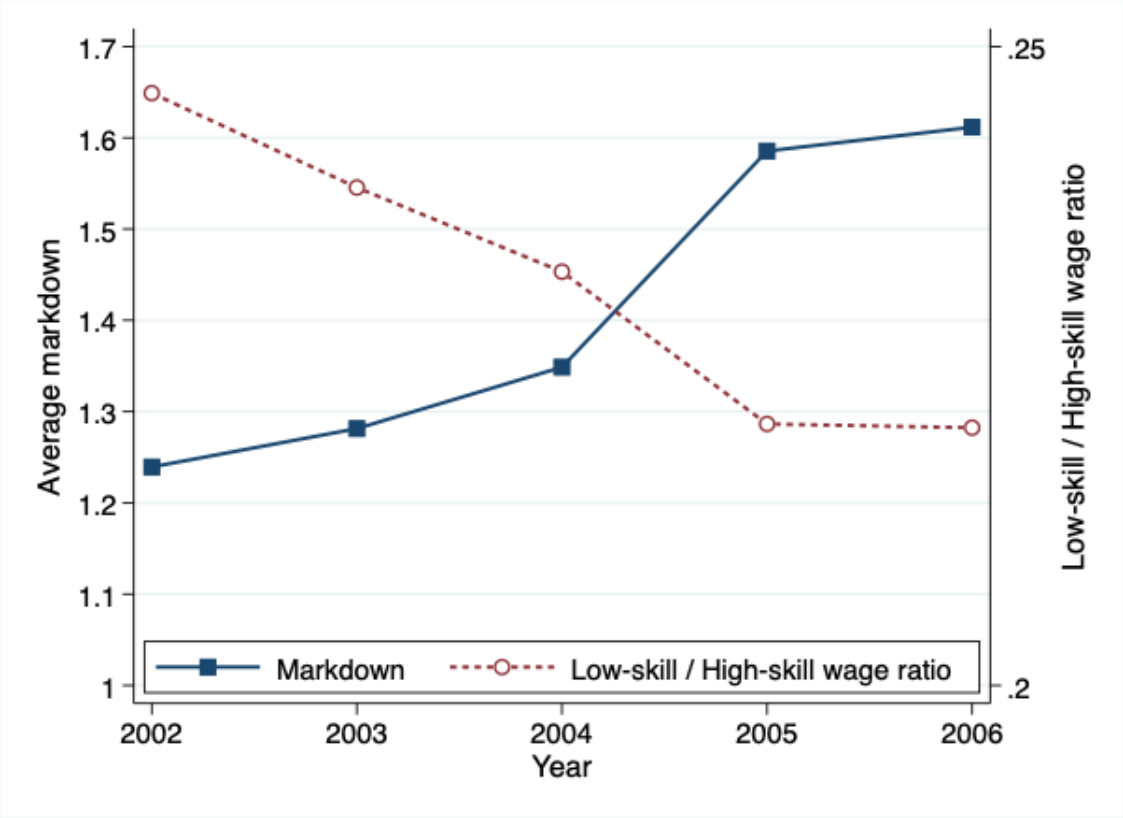


Figure A.1 Average markdown and the average wage ratio between low-skilled and high-skilled workers, 2002–2006

Notes: This graph plots the time path of average markdowns of Chinese firms and the average relative wage of low-skilled workers to high-skilled workers over 2002–2006. Markdowns are computed as *output-weighted* averages at the year level. The average relative wage is defined as the ratio of the average wage of high-skilled workers (those with an associate degree or higher) to low-skilled workers, with wages deflated using province-level CPI indices. Due to the 2002 redefinition of the UHS dataset, where the survey population changed from nationwide non-agricultural households (before 2002) to urban residents in cities and county-level urban areas (after 2002), data from 2002 onwards are used to ensure comparability across years.
Data sources: ASIF (Annual Survey of Industrial Firms) and UHS (Urban Household Survey).

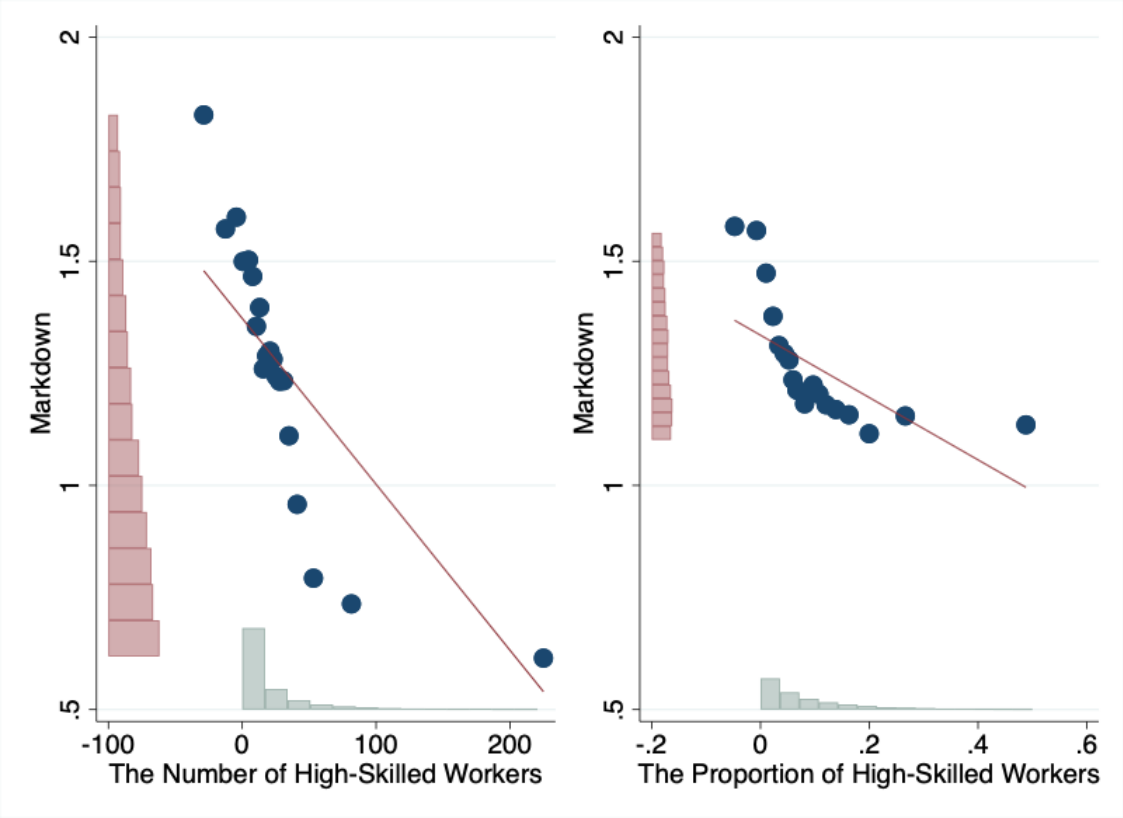


Figure A.2 The cross-sectional relationship between markdown and skill intensity

Notes: This figure plots the cross-sectional relationship between markdown and skill intensity, using both the *level* and the *share* of high-skilled workers. Both specifications absorb province fixed effects and industry fixed effects, and figures are constructed using binned scatter plots of the residualized variables.
Data sources: the 2004 ASIF (Annual Survey of Industrial Firms) and the 2004 Economic Censuses. High-skilled workers are defined as those with an associate degree or above. The proportion of high-skilled workers is the share of high-skilled employment in total employment.

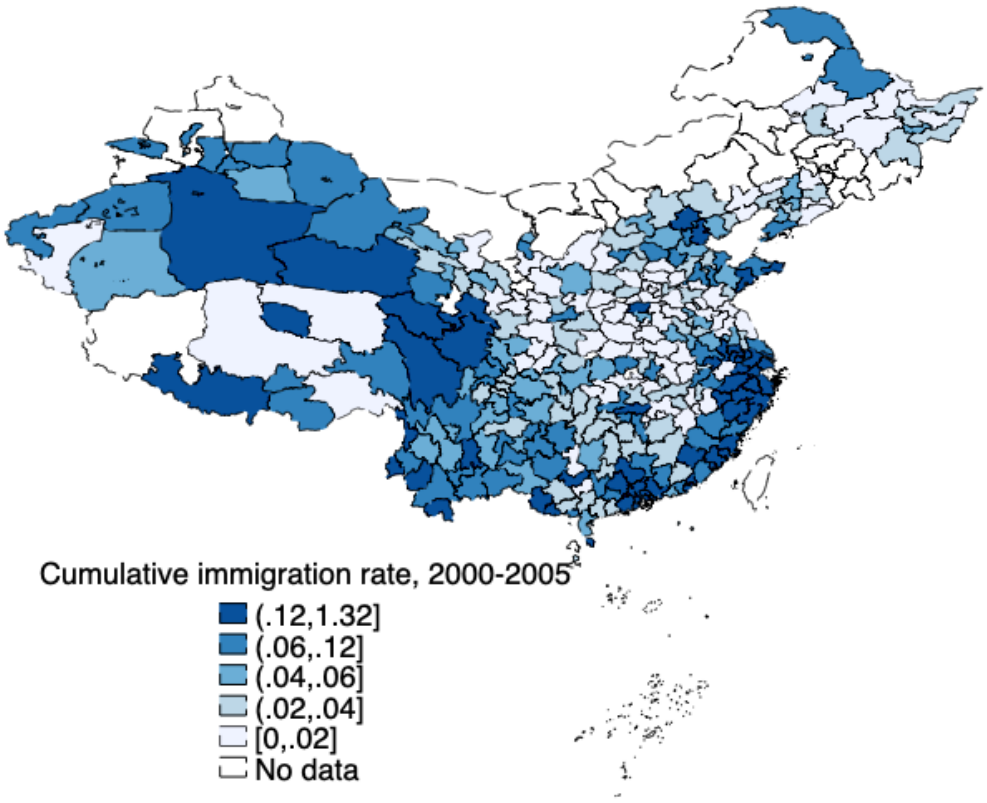


Figure A.3 Immigration Rate Across Chinese Prefectures (2000-2005)

Notes: This figure shows the cumulative immigration rate across Chinese prefectures between 2000 and 2005.
Data sources: 1 Percent Population Survey of 2005 by NBS.

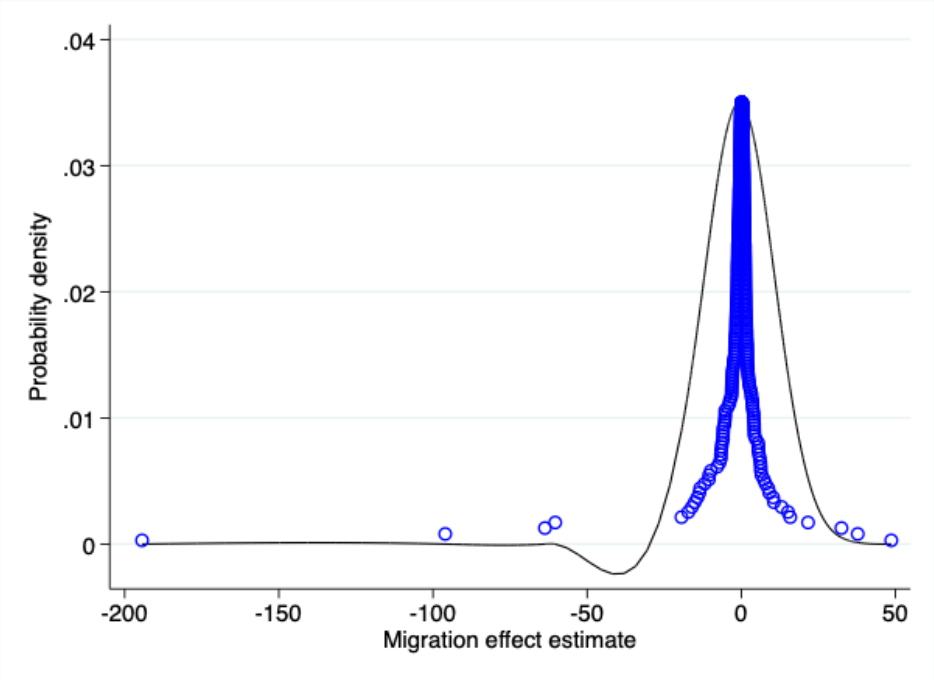


Figure A.4 Distribution of Estimated Coefficients in the Placebo Test

Note: The figure shows the distribution of the estimated coefficients from 500 simulations randomly assigning migration supply shocks to cities.

B Empirical Appendix

B.1 Data Appendix

Chinese Manufacturing Firms Dataset Our main firm-level dataset is the Annual Survey of Industrial Firms (ASIF) provided by China's National Bureau of Statistics (NBS). We conduct the following data cleaning process. First, we exclude observations with missing values in key variables, such as total assets, employment, sales, the annual average book value of fixed assets, and gross output value. Second, we enforce the above-scale criterion consistently by removing observations that do not meet the threshold (for example, non–state-owned firms with annual sales below RMB 5 million). Third, we drop observations that violate standard accounting rules , such as total assets being smaller than current assets, smaller than net fixed assets, smaller than fixed assets, or accumulated depreciation being smaller than current-period depreciation. Finally, we address extreme values by winsorizing key variables (e.g., employment and gross output) at 1% level.

After these steps, we harmonize industry codes across years to a consistent four-digit manufacturing classification to account for the 2002 revision of China's Industrial Classification (CIC); see Brandt et al. (2017). Our analysis focuses on manufacturing firms. In the main regressions, following Imbert et al. (2022), we use a balanced panel of firms observed in every year from 2000 to 2006, comprising 29,616 firms.

Migrants and Agriculture Datasets We borrow the construction of migration flows from Imbert et al. (2022). Specifically, we use the nationally representative 2005 1% Population Survey ("Mini-Census") administered by the NBS, which enumerates individuals at their current place of residence—migrants included—and records hukou type (agricultural vs. non-agricultural), place of registration and residence (prefecture level), occupation, industry, income, education, and housing. Crucially for flows, the survey asks the main reason for leaving the place of registration, the year of departure, and place of residence one and five years prior to the interview. Following Imbert et al. (2022), we combine these items to build an origin–destination–year matrix of rural-to-urban flows between all prefecture pairs for 2000–2005; we restrict to migrants aged 15–64 at the time of migration and exclude those migrating for study (less than 5% of total). We adjust flows for return migration (which is non-negligible), while step migration is found to be negligible; details follow the original implementation. In addition, to measure baseline settlement patterns (used as shares in the shift–share design), we rely on a 1% extract of the 2000 Population Census (NBS, 2000) and apply the same migrant definition as in the 2005 Mini-Census.

This procedure yields annual emigration flow by origin and immigration flow by destination that are comparable across space and time and suitable for merging to firm outcomes at the prefecture level.

The agricultural shocks used for the shift–share instrument follow Imbert et al. (2022). Baseline cropping portfolios at the prefecture level are measured from a geo-coded map of harvested areas in China from the 2000 World Census of Agriculture, which we combine with potential yields from the FAO/IIASA Global Agro-Ecological Zones (GAEZ). For each prefecture–crop cell, baseline agricultural potential is the product of harvested area and GAEZ potential yield; these portfolios are normalized to shares and held fixed over our sample window. International crop prices are taken from FAO Agricultural Producer Prices (with robustness to the World Bank "Pink Sheet" series). For each major crop, we construct a world price series excluding China to avoid domestic endogeneity. Origin–year agricultural shocks are then formed by interacting the time-invariant prefecture portfolios with the external (world) price movements for the corresponding crops.

Firm Census Datasets: To construct changes in firm-level labor composition, we supplement ASIF with information on employment by education from two censuses. First, the *China Industrial Census 1995* (NBS) enumerates all industrial enterprises and production units nationwide and reports, at the legal-entity level, basic identifiers and registry information, location and industry codes, and employment by education. Second, the *China Economic Census 2004* (NBS) updates firm registers and reporting standards and extends coverage to the entire secondary and tertiary sectors. It also records enterprise identifiers, industry and location, and employment disaggregated by education.

We extract, for each census year, the number of employees by education category and map these to firms used in our ASIF panel data.

Urban Household Survey Dataset To estimate labor-supply elasticities by skill, we complement firm and migration data with the Urban Household Survey (UHS) administered by the Urban Social and Economic Survey Division of China's National Bureau of Statistics (NBS). The UHS is a large, nationally coordinated household survey of urban residents that reports detailed household demographics and composition, labor-market status, cash income and expenditures, quantities and values of major commodity purchases, housing conditions, and ownership of durable goods.

The UHS is a diary-based household survey designed and fielded by NBS Urban Socio-Economic Survey Teams using stratified, multi-stage probability sampling with probability proportional to size. The sampling design follows a rotating panel structure in which fixed

sample households keep continuous records; each panel lasts three years with one third of households refreshed annually. During the period from 2002 to 2006, the national sample typically covers 16 provinces [13] and about 25,000 centrally sampled households per year (approximately 50,000 including locally sampled households drawn under the same protocol). In our analysis, we use UHS individual records to construct wage measures by education-based skill group.

Summary Statistics: Table B.1 reports summary statistics for the main variables used in the paper.

Table B.1 Summary statistics of key variables

Variable	Definition	Obs.	Mean	Std. dev.	Min	Max
Prefecture level						
Migrants	Sum of yearly migration flows over the period, divided by the baseline non-migrant population	288	0.111	0.201	0.003	1.323
IV	Rural income shocks at migrant-sending regions, weighted by historical migration patterns	288	0.138	0.058	0.001	0.259
Firm level						
Δ markdown	Change in markdown between 2006 and 2000	29,616	0.666	1.583	-10.417	11.116
Δ skill intensity	Change in skill intensity between 2004 and 1995	36,061	0.070	0.141	-1	1

Notes: City-level variables are constructed following Imbert et al. (2022). Firm-level change in markdown is computed from ASIF and our estimation. Firm-level change in skill intensity is constructed from the 1995 Industrial Census and the 2004 Economic Census.

[13] The 16 provinces covered by the UHS include Beijing, Shanxi, Liaoning, Heilongjiang, Shanghai, Jiangsu, Anhui, Jiangxi, Shandong, Henan, Hubei, Guangdong, Chongqing, Sichuan, Yunnan, and Gansu.

B.2 Markdown Estimation

Definition of Wage Markdown In a perfectly competitive labor market, firms are wage takers, facing a perfectly elastic labor supply. Wages are pinned down by market forces and equal the MRPL, leaving no wedge between labor's contribution and compensation. However, real-world labor markets frequently display frictions, such as limited worker mobility, information asymmetry, and institutional constraints, which grant employers wage-setting power (Manning, 2003b; Alan, 2011; Ashenfelter et al., 2010). Consider a simple theoretical framework where firms choose labor input L to maximize profits:

$$\max_{L} \ \pi(L) = R(L) - w(L) \cdot L,$$

where $R(L)$ is revenue and $w(L)$ is the upward-sloping wage schedule. The firm's optimality condition implies:

$$MRPL = \frac{\partial R}{\partial L} = \frac{\partial [w(L) \cdot L]}{\partial L} = MWCL,$$

where $MWCL$ denotes the marginal wage cost. Because hiring additional workers requires raising wages not only for new but also for existing employees, the marginal cost of labor exceeds the wage:

$$MWCL = w(L) \left(1 + \frac{\partial w}{\partial L} \cdot \frac{w}{L}\right) = w(L) \cdot \chi_L,$$

where $\chi_L > 1$ captures the labor market wedge associated with monopsony. This wedge creates a gap between the marginal product of labor (MRPL) and the actual wage paid to the worker. Hence, markdown is defined as:

$$markdown_{it} \equiv \frac{MRPL_{it}}{Wage_{it}} \tag{B.1}$$

which directly measures the extent to which firms extract labor surplus. Under perfect competition, $\chi_L = 1$ implies markdown $= 1$. Under monopsony power, the markdown exceeds one, indicating that workers receive less than their marginal product and firms capture the residual rents.

We follow the now-standard production-based approach of (De Loecker and Warzynski, 2012) (DLW), as implemented in recent studies on labor market power (e.g., Yeh et al. (2022)). The method combines production function estimation with the flexible-input condition to recover firm–year output elasticities and markups, which we then map into markdowns under a wage-setting framework with finite labor supply elasticity.

Production function estimation Let $y_{it} \equiv \ln Q_{it}$ denote (log) output produced with logged inputs of labor ℓ_{it}, capital k_{it}, and materials m_{it}. We estimate a flexible translog technology with a productivity term ω_{it} and i.i.d. error ε_{it}:

$$y_{it} \;=\; f(\ell_{it}, k_{it}, m_{it}; \beta) \;+\; \omega_{it} \;+\; \varepsilon_{it}, \tag{B.2}$$

where the translog mapping $f(\cdot)$ is *explicitly* given by

$$\begin{aligned} f(\ell, k, m; \beta) = {} & \beta_\ell \ell_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{\ell\ell} \ell_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 \\ & + \beta_{\ell k}\, \ell_{it} k_{it} + \beta_{\ell m}\, \ell_{it} m_{it} + \beta_{km}\, k_{it} m_{it} + \beta_{\ell km}\, \ell_{it} k_{it} m_{it}. \end{aligned} \tag{B.3}$$

All variables entering (B.3) are in *real terms* (deflated revenue and deflated inputs; see Data Appendix). Translog nests Cobb–Douglas and allows output elasticities to vary with input intensities, avoiding confounding technological heterogeneity with markups.

We estimate the production function industry-by-industry (at the 2-digit level) and treat capital as a dynamic (state) input while allowing materials and labor to be static/flexible inputs. This structure follows the ACF approach adapted by De Loecker and Warzynski (2012).

To address simultaneity between input choices and productivity, we proxy ω_{it} with a (strictly) increasing material demand function.

$$m_{it} = m_t\big(k_{it}, \omega_{it}, z_{it}\big) \quad \Rightarrow \quad \omega_{it} = h_t\big(m_{it}, k_{it}, z_{it}\big),$$

where z_{it} collects observed shifters of input demand (e.g., export status e_{it}; ownership). Substituting the proxy into (B.2) gives the *first stage*:

$$\begin{aligned} y_{it} &= \phi_t\big(\ell_{it}, k_{it}, m_{it}, z_{it}\big) + \varepsilon_{it}, \\ \phi_t(\cdot) &\equiv f(\ell_{it}, k_{it}, m_{it}; \beta) + h_t\big(m_{it}, k_{it}, z_{it}\big). \end{aligned} \tag{B.4}$$

from which we obtain estimates of expected output, $(\hat{\phi}_{it})$, and the residual $\hat{\varepsilon}_{it}$.

Specifically, we approximate $h_t(\cdot)$ flexibly and run the first stage regression by including (i) the translog $f(\ell_{it}, k_{it}, m_{it}; \beta)$, (ii) interactions of the input terms in $f(\cdot)$ with the export-status indicator e_{it} to allow the proxy to shift with exporting, and (iii) a rich set of fixed effects, including ownership, year, province, and 4-digit industry, absorbed in the regression.

After the first-stage estimation, we use the predicted value $\hat{\phi}_{it}$ to construct productivity for any parameter vector $\beta = (\beta_\ell, \beta_k, \beta_m, \beta_{\ell\ell}, \beta_{kk}, \beta_{mm}, \beta_{\ell k}, \beta_{\ell m}, \beta_{km}, \beta_{\ell km})$ as follows:

$$\begin{aligned} \widehat{\omega}_{it}(\beta) = {} & \hat{\phi}_{it} - \beta_\ell \ell_{it} - \beta_k k_{it} - \beta_m m_{it} - \beta_{\ell\ell} \ell_{it}^2 - \beta_{kk} k_{it}^2 - \beta_{mm} m_{it}^2 \\ & - \beta_{\ell k} \ell_{it} k_{it} - \beta_{\ell m} \ell_{it} m_{it} - \beta_{km} k_{it} m_{it} - \beta_{\ell km} \ell_{it} k_{it} m_{it}. \end{aligned} \tag{B.5}$$

In the second stage, we proxy the law of motion of productivity ω_{it} as follows:

$$\omega_{it} = g\big(\omega_{i,t-1}\big) + \xi_{it} \tag{B.6}$$

and obtain the estimate of the residual $\widehat{\xi}_{it}(\beta)$ by regressing $\widehat{\omega}_{it}(\beta)$ on its lag.

We then use the following moment conditions to estimate parameters in the production function.

$$\mathbb{E}\left[\xi_{it}(\beta)\begin{pmatrix} \ell_{i,t-1} \\ k_{i,t} \\ m_{i,t-1} \\ \ell_{i,t-1}^2 \\ k_{i,t}^2 \\ m_{i,t-1}^2 \\ \ell_{i,t-1}k_{i,t} \\ \ell_{i,t-1}m_{i,t-1} \\ k_{i,t}m_{i,t-1} \\ \ell_{i,t-1}k_{i,t}m_{i,t-1} \end{pmatrix}\right] = \mathbf{0}. \tag{B.7}$$

We estimate β by GMM using (B.7), initialize at OLS.

Markup estimation. Under the translog specification in (B.3), the (firm–year) output elasticities of labor and materials are

$$\theta_{it}^{\ell} \equiv \frac{\partial f}{\partial \ell} = \beta_{\ell} + 2\beta_{\ell\ell}\,\ell_{it} + \beta_{\ell k}\,k_{it} + \beta_{\ell m}\,m_{it} \;+\; \beta_{\ell km}\,k_{it}m_{it}, \tag{B.8}$$

$$\theta_{it}^{m} \equiv \frac{\partial f}{\partial m} = \beta_{m} + 2\beta_{mm}\,m_{it} + \beta_{\ell m}\,\ell_{it} + \beta_{km}\,k_{it} \;+\; \beta_{\ell km}\,\ell_{it}k_{it}. \tag{B.9}$$

Taking materials as the flexible input, the static cost-minimization FOC implies the familiar ratio of an input's elasticity to its revenue share:

$$\mu_{it} \;=\; \frac{\theta_{it}^{m}}{\alpha_{it}^{m}}, \tag{B.10}$$

where α_{it}^{m} is the observed revenue share of materials. Since the production function is estimated on (deflated) revenue rather than physical output, we strip out the idiosyncratic error using the first-stage residual $\widehat{\varepsilon}_{it}$ and compute

$$\alpha_{it}^{m} \;=\; \frac{\text{materials expenditure}_{it}}{\text{revenue}_{it}/\exp(\widehat{\varepsilon}_{it})}. \tag{B.11}$$

Equations (B.10)–(B.11) deliver firm–year markups once θ_{it}^{m} is constructed from the estimated translog parameters.

Markdown estimation. Let wages be set along a differentiable, finitely elastic firm–level labor supply schedule. The markdown ν_{it} is defined as the ratio of the marginal revenue product of labor (MRPL) to the wage paid:

$$\nu_{it} \equiv \frac{MRPL_{it}}{W_{it}}. \tag{B.12}$$

Using the link between markups and revenue shares, and under the wage–setting assumptions in (Yeh et al., 2022), the markdown can be expressed in terms of observable shares and estimated elasticities as

$$\nu_{it} = \frac{\theta_{it}^{\ell}}{\alpha_{it}^{\ell}} \cdot \mu_{it}^{-1}, \tag{B.13}$$

where θ_{it}^{ℓ} is the output elasticity of labor and α_{it}^{ℓ} is labor's revenue share. In the competitive benchmark, $\nu_{it} = 1$. Values above unity indicate that firms pay wages below MRPL and thus exercise monopsony power.

Table B.2 reports the descriptive statistics for firm markdowns by industry. For each two-digit CIC industry, we list the number of firm-year observations and the mean markdown.

Table B.2 Descriptive statistics of firm markdown by industry

Code	Industry name	Observations	Mean
13	Processing of Agricultural and Sideline Food Products	13,824	3.327
14	Food Manufacturing	7,214	2.325
15	Beverage Manufacturing	4,853	2.260
16	Tobacco Products	325	2.874
17	Textile Industry	17,267	0.612
18	Manufacture of Textile Apparel, Footwear and Headgear	13,362	1.147
19	Leather, Fur and Feathers and Related Products	6,606	1.310
20	Wood Processing and Products of Wood, Bamboo, Rattan, Palm and Straw	3,337	1.211
21	Furniture Manufacturing	2,098	0.621
22	Paper and Paper Products	8,316	0.699
23	Printing and Reproduction of Recorded Media	7,248	1.335
24	Manufacture of Cultural, Educational and Sports Goods	3,785	0.463
25	Petroleum Processing, Coking and Processing of Nuclear Fuel	1,160	0.926
26	Manufacture of Chemical Raw Materials and Chemical Products	17,049	0.937
27	Pharmaceutical Manufacturing	7,720	3.393
28	Manufacture of Chemical Fibers	1,125	1.093
29	Rubber Products	2,999	0.662
30	Plastics Products	12,549	1.507
31	Non-metallic Mineral Products	22,219	0.962
32	Smelting and Rolling of Ferrous Metals	3,548	0.986
33	Smelting and Rolling of Non-ferrous Metals	5,055	4.204
34	Metal Products	11,626	0.845
35	General Equipment Manufacturing	21,053	0.993
36	Specialized Equipment Manufacturing	12,808	2.136
37	Transport Equipment Manufacturing	13,512	1.094
39	Electrical Machinery and Equipment Manufacturing	14,278	0.735
40	Manufacture of Communication Equipment, Computers and Other Electronic Equipment	9,051	0.729
41	Instruments and Meters; Cultural and Office Machinery Manufacturing	4,056	1.025
42	Other Manufacturing	6,746	1.455

Notes: Observations count firm-year records within each two-digit industry; "Mean" is the industry average of firm markdown.

B.3 Firm-level Labor Supply Elasticity Estimation

Following Rubens et al. (2024), we adopt a discrete-choice nested logit model of labor supply in the tradition of Berry (1994), which has been applied in labor market contexts by Berger et al. (2022). In this framework, manufacturing workers i in labor market ℓ choose among a set of firms within that market. Labor markets are defined at the prefectural city level, while nests are constructed at the county-by-4-digit-industry level, allowing for worker mobility both across industries and counties. In addition, workers may exit the non-farm manufacturing (NFM) sector and opt for an outside alternative $f = 0$, which is modeled as a separate nest.

Each worker's utility from being employed by firm f at time t is given by:

$$U_{jft} = \underbrace{\gamma_t W^l_{ft} + \gamma_X X_{ft} + \xi_{ft}}_{\equiv \delta_{ft}} + \sum_n (d_{fn}\zeta_{jn}) + (1-\varsigma)\varepsilon_{jft} \tag{B.14}$$

where W^l_{ft} is the wage, X_{ft} denotes observed firm characteristics, ξ_{ft} captures unobserved firm amenities, ζ_{jn} is a nest-level random taste shock, and ε_{jft} is an idiosyncratic shock. The parameter γ_t reflects workers' marginal utility of wages and is allowed to vary over time to capture evolving labor market conditions. We model this time variation linearly: $\gamma_t = \gamma_0 + \gamma_1 t$.

Normalizing the utility of the outside option to zero, the labor market share for firm f is:

$$s_{ft} - s_{0t} = \gamma_t W^l_{ft} + \varsigma s^n_{ft} + \gamma_X X_{ft} + \xi_{ft} \tag{B.15}$$

where s^n_{ft} is the log share of firm f within its nest.

To estimate Equation (B.15), we instrument for wages and within-nest market shares, since wages may be endogenously determined by unobserved amenities ξ_{ft}. We employ two instruments. First, we use local migration inflows in year t, which plausibly shift labor supply and hence firms' wage-setting behavior. Second, we use the number of firms within each industry-year-county cell, which provides exogenous variation in market structure relevant for identifying the nesting parameter ς.

The outside option is measured as the total city population minus total manufacturing employment. Labor market shares at both the total market and nest level are calculated using firm-level employment data. The vector X_{ft} includes sector fixed effects to control for time-invariant heterogeneity across industries, city fixed effects to account for persistent regional differences in labor market conditions, ownership type indicators (e.g., whether the firm is foreign-invested) to capture variation in non-wage amenities across ownership structures, and year fixed effects to absorb macroeconomic shocks and national policy changes over time.

Using the estimated parameters ς and γ_t, we compute each firm's inverse labor supply elasticity ψ^l_{ft} as follows.

$$\psi^{\ell}_{ft} = 1 + \frac{1-\varsigma}{\gamma_t W^{\ell}_{ft}\big(1-\varsigma s^n_{ft} - (1-\varsigma)s_{ft}\big)}\,. \tag{B.16}$$

The estimated elasticity captures how sensitive a firm's labor supply is to changes in wages and enables us to investigate whether migrant inflows alter the degree of firms' monopsony power in local labor markets.

C Theory Appendix

Derivation of migrants' labor supply (3). Consider a rural resident i deciding whether to migrate to an urban manufacturing firm. If she chooses to work in sector s and firm f, her indirect utility is:

$$V_{isf} = \ln\left(\frac{w^*_{sf}}{w^r d}\right) + \epsilon_{isf}, \qquad s \in S,\ f \in N_s, \tag{C.1}$$

where w^*_{sf} is the wage paid by firm f in sector s; w^r is the agricultural income (opportunity cost of migration); d is the migration cost (transportation, psychological, etc.); ϵ_{isf} is an idiosyncratic preference shock.

The preference shocks $\{\epsilon_{isf}\}$ are assumed to follow a generalized extreme value (GEV) distribution:

$$G_\epsilon(\epsilon) = \exp\left\{ -\left[\sum_{s\in S}\left(\sum_{f\in N_s}\exp\left(-\frac{\epsilon_{isf}}{\kappa_s}\right)\right)^{\frac{\kappa_s}{\kappa}}\right]^{\kappa}\right\}, \tag{C.2}$$

where $\kappa_s \in (0,1)$ is the within-sector substitution parameter; $\kappa > 0$ is the between-sector substitution parameter; S is the set of manufacturing sectors; N_s is the set of firms in sector s.

The choice structure consists of two levels. At the upper level, the farmer chooses a sector s. At the lower level, given the sector s, she chooses a firm f within that sector.

The inclusive value of sector s is:

$$I_s = \ln\left(\sum_{f\in N_s}\exp\left(\frac{\ln(w_{sf}/w^r d)}{\kappa_s}\right)\right) = \ln\left(\sum_{f\in N_s}\left(\frac{w_{sf}}{w^r d}\right)^{\frac{1}{\kappa_s}}\right). \tag{C.3}$$

The probability of choosing sector s is:

$$P_s = \frac{\exp\left(\frac{\kappa_s}{\kappa} I_s\right)}{\sum_{s'\in S}\exp\left(\frac{\kappa_{s'}}{\kappa} I_{s'}\right)}. \tag{C.4}$$

The conditional probability of firm f given sector s is:

$$P_{f|s} = \frac{\exp\left(\frac{\ln(w_{sf}/w^r d)}{\kappa_s}\right)}{\sum_{f'\in N_s}\exp\left(\frac{\ln(w_{sf'}/w^r d)}{\kappa_s}\right)} = \frac{(w_{sf}/w^r d)^{\frac{1}{\kappa_s}}}{\sum_{f'\in N_s}(w_{sf'}/w^r d)^{\frac{1}{\kappa_s}}}. \tag{C.5}$$

Then the joint probability is:

$$P_{sf} = P_s \cdot P_{f|s}. \tag{C.6}$$

The migrant labor supply to a specific firm f in sector s can be written as:

$$q_{sf} = \bar{L}^r \cdot P_{sf} = \bar{L}^r \cdot P_s \cdot P_{f|s}. \tag{C.7}$$

Substituting equations (C.4) and (C.5):

$$q_{sf} = \bar{L}^r \cdot P_s \cdot \frac{(w_{sf}/w^r d)^{\frac{1}{\kappa_s}}}{\sum_{f' \in N_s} (w_{sf'}/w^r d)^{\frac{1}{\kappa_s}}}. \tag{C.8}$$

Define the sector-specific constant:

$$\lambda_s \equiv P_s \cdot \frac{1}{\sum_{f' \in N_s} (w_{sf'}/w^r d)^{\frac{1}{\kappa_s}}}. \tag{C.9}$$

Then the migrant labor supply to firm f can be written as:

$$q_s f = \left(\bar{L}^r \lambda_s\right) \cdot \left(\frac{w^*_{sf}}{w^r d}\right)^{\frac{1}{\kappa_s}}. \tag{C.10}$$