# The Middlemen of Development?

## Evidence from Labor Recruitment Agencies in China

Linchuan Xu *

May 31, 2026

[latest version]

**Abstract**

Can profit-driven middlemen arise endogenously to correct aggregate misallocations and promote development? This paper studies the role of recruitment agencies in facilitating spatial labor reallocation in China. Using a newly constructed dataset, I find that such profit-driven labor intermediaries established branches to arbitrage away pre-existing inter-regional wage disparities. Exploiting the staggered rollout of agencies in difference-in-differences and IV designs, the analysis shows that the expansion of agency networks facilitated temporary mobility of low-skilled workers and reduced regional wage dispersion by 10%. This translates into 3% aggregate output gains by 2010, of which agencies appropriated 40% as profits. I develop a neoclassical growth model with endogenous agency dynamics, where initial misallocations create pecuniary incentives for agencies who in turn speed up regional convergence. Promoting agency workers' transition into the permanent labor force in receiving regions can generate large efficiency gains while also ensuring these gains are more widely shared.

**JEL Classification:** O1, O4, J6, E2

**Keywords:** Misallocations, Agencies, Mobility, Convergence

*Department of Economics, London School of Economics and Political Science. Email: l.xu25@lse.ac.uk. I am grateful for Timothy Besley, Maitreesh Ghatak, and Alwyn Young for their continuous guidance and encouragement on this paper. I also thank Britta Augsburg, Clare Balboni, Robin Burgess, Mike Callen, Matthias Doepke, Dave Donaldson, Alan Manning, Andrew Newman, Rachel Ngai, Trinh Pham, Steve Pischke, Imran Rasul, Christiane Szerman, John Van Reenen, Martin Rotemberg, Yanos Zylberberg, and seminar participants at LSE, Manchester, Bristol, IFS, NEUDC 2025, and EWMES 2025 for helpful comments and suggestions. I also thank Lingyun Zhang, Tianyu Zhang, and Liuxian Li for valuable research assistance. I acknowledge financial support from STICERD at LSE. This paper was previously circulated with the title "Profit from Wedges: Evidence from Labor Recruitment Agencies in China".

# 1 Introduction

A central question in the misallocation literature, to extend Lucas (1990)'s famous question about international capital flow, is why *all* factors of production do not automatically flow to their most productive uses. Recent papers following Hsieh and Klenow (2009) have shown that promoting factor reallocation by removing the distortionary "wedges" can, at least potentially, generate substantial productivity and output gains in many empirical contexts. But a natural Schumpeterian question follows[1]: if the potential gains are so large, wouldn't one expect profit driven economic agents, i.e., entrepreneurs or market intermediaries, to arise endogenously and deliver them? On the other hand, a large body of literature on market intermediaries suggests that middlemen of various forms do arise endogenously to address market frictions (Gong et al., 2025). But are they powerful enough to meaningfully generate aggregate gains and shape the development process?

Empirically, while many developing economies, such as those in Sub-Saharan Africa, have experienced widening wage disparities along their development paths, wage dispersion across firms and regions in China *reduced* by 46% in the manufacturing sector over 1999–2015. China's reforms on the land market, privatization, and trade openness have been extensively studied as drivers of its growth experience (for example, Lin, 1992; Lau et al., 2000; Young, 2000). Large-scale changes in the *labor* market, by contrast, have received much less attention. This is most notable for labor recruitment agencies: legally recognized by the 2008 Labor Contract Law, these agencies are widely believed to employ a large fraction of the workforce, yet they have been largely overlooked in the economics literature. Whether and how these agencies contributed to China's wage convergence and regional development remains an open question.

This paper studies the expansion of labor recruitment agencies in the Chinese context well known for spatial misallocations of labor arising from informational and search-related frictions[2]. Through their local branches, agencies channel surplus labor from across the country into auxiliary positions in distant workplaces[3]. This has helped meet the growing demand for low-skilled workers in, for example, coastal cities—a task that became increasingly pertinent as China entered its labor-intensive, export-led growth episode in the early 2000s (e.g., State Council, 2006). Such reallocation in turn narrows wage gaps across regions and partially mitigates the initial misallocations. To study this process, I construct a comprehensive dataset covering the near-universe of labor recruitment agencies in China[4].

[1] By definition in Schumpeter (1934), entrepreneurs "have not accumulated any kinds of goods, they have created no original means of production, but have employed existing means of production differently, more appropriately, more advantageously. They have 'carried out new combinations.' " He also argued that the surplus that is generated should fall onto the entrepreneurs as profits since they are the one "whose achievement is to introduce the looms" (page 132).

[2] There are many sources of labor misallocations in China, most notably the *Hukou* registration system which requires migrants to seek approval from local government to change their officially registered location (Ngai et al., 2019; Gai et al., 2025). Furthermore, there used to be official requirements for migrants to work across regions, but these have been largely relaxed (Ebenstein and Zhao, 2015): for example, the process of obtaining a temporary residence permit has been streamlined in 2003. The major obstacle for workers to live in a city without *hukou* is associated with access to public services, healthcare, and educational resources (Meng, 2012; Tombe and Zhu, 2019). But agencies are less likely to help workers to overcome the *hukou* barrier: labor agencies might send workers to multiple places without changing their *Hukou* status. Second, agency workers are mostly less educated and low-skilled, and *Hukou* relaxation for this segment of the workforce is rare. In general it is rare for temporary low-skilled migrants to obtain *hukou* in richer cities (Meng, 2012).

[3] For example, Baseler (2023) showed that urban earning are underestimated in Kenya and providing better information can increase migration to the capital city. In this paper, labor agencies can provide information on specific job opportunities to workers and facilitate their travels and accommodation (see the case studies by China Labor Watch (2016)).

[4] They are formally known as Labor Dispatch Agencies (LDA). See Yan (2023) for a related paper that documents the rise of LDA in China. In practice, it is very difficult to distinguish labor dispatch from labor outsourcing in the Chinese context, as the two contractual arrangements are often used interchangeably and the legal boundary between them remains poorly defined (see Section 2.1). Rather than distinguishing between these specific contractual forms, this

Using this novel dataset, the paper quantifies how the spatial expansion of these agencies contributed to wage convergence through inter-regional labor reallocation, generating aggregate efficiency gains while capturing part of the resulting surplus as profits.

The core argument of this paper is that factor misallocations are not merely a static distortion to be removed, but an endogenous object that creates its own corrective force: the regional wage disparities generated by mobility frictions are precisely the arbitrage opportunities that draw profit-driven intermediaries into existence. Most of the misallocation literature treats the distortionary "wedges" as exogenous, policy-like primitives (for example, Sraer and Thesmar, 2023; Bau and Matray, 2023); here, by contrast, the wedge calls forth the middlemen that erode it, which then generate aggregate impact. Second, this erosion is inherently *dynamic*. I take a longer-run perspective and examine the role of labor market intermediaries along the transition path of regional convergence à la Solow (1956). Here, the misallocation wedge is a transitional phenomenon, arbitraged away over time by endogenous agency dynamics. Third and importantly, the *distribution* of the surplus recovered from reducing misallocations is shaped by the market structure of the intermediaries, and this distribution carries efficiency consequences. I show that labor market intermediaries capture a substantial share of these gains as profits, and that greater competition among them amplifies efficiency both directly, by strengthening incentives for labor reallocation, and indirectly, by shifting a larger share of the surplus to producers and thereby accelerating capital accumulation.

I first show empirically that the location choice of privately owned labor recruitment agencies exhibits a U-shaped relationship with pre-existing average local wage, whereas no such pattern is observed for state-owned agencies. In other words, they establish branches to connect cities at the two extremes of the pre-existing regional wage distribution, reallocating labor from the least developed areas to high wage cities where they are needed the most. Furthermore, drawing on a sub-sample of labor agencies from the administrative tax survey dataset (2007-2016), I find a strong positive correlation between pre-existing wage disparity and their annual profits, and the number of competing agency networks reduces profits. This echoes with the idea that among cities where labor misallocations are the greatest, there are also greater (pecuniary) incentives for intermediaries to establish branches there to exploit the arbitrage opportunities.

This paper then provides empirical evidence on how the expansion of agencies networks affects factor reallocation and regional convergence. Using population census data in 2000, 2005 and 2010 and exploiting the staggered rollout of agencies, I find that when agencies established branches between cities, the temporary flow of workers (i.e., those that did not change their *hukou* status) also increased. In particular, labor flows increase more when origin and destination cities have larger pre-existing wage gaps, implying that agency networks move workers from low to high wage cities. Furthermore, the expanding agency networks have a stronger effect on temporary flow of below college, low-skilled workers than for workers with at least a college degree, which is consistent with the role of labor agencies in organizing predominantly low-skilled workers. These patterns of labor reallocation, in turn, contribute to regional convergence in wages and output per worker, consistent with the theoretical prediction. With a sample of manufacturing firms in the tax survey dataset, I calculate average wage for each city and wage gaps for all pairs of cities. A 1% larger initial wage disparity is associated with 0.28% additional subsequent reduction of the wage gap once cities are connected by agency branches. Moreover, the extensive margin of agency connection among cities matters quantitatively more for reducing wage and output per worker gaps than the intensive margins, i.e., the number of

paper focuses on these labor intermediaries themselves and studies how they promote inter-regional labor reallocation through their spatial branch networks.

agency groups that connect the same pairs of cities. I interpret this as evidence that while competition among agency networks may bring down their profits, it may be less important for the arbitrage forces to make the law of one price to hold. Besides, using event study estimators proposed by Callaway and Sant'Anna (2021), de Chaisemartin and D'Haultfoeuille (2024) as well as Arkhangelsky et al. (2021), I find larger effect of wage gap reduction in later years after agencies build up their linkages across cities, suggesting that these cities embark on the dynamic path of wage convergence.

A few additional pieces of evidence support the labor flow and wage convergence results. Firstly, to complement the standard difference-in-differences framework, I leverage exogenous variation generated by the 2013 Amendment to the Labor Contract Law, using city pairs' pre-reform exposure to the policy shock as an instrument for agencies expansion across cities. The IV estimates reveal increased temporary labor flows and a narrowing of regional wage gaps, broadly consistent with the baseline DiD results. I further control for road connectivity across city pairs to mitigate concerns that improvements in transportation infrastructure may bias the estimates. Secondly, the expansion of labor agencies leads to similar results on wage convergence in sectors such as construction, hospitality, and logistics, which are similarly likely to employ agency workers. Thirdly, using China Labor Dynamics Survey (2012-2018), I calculate average city wages separately for workers with and without a college degree. I find a stronger effect on regional wage convergence among less-skilled workers, which resonates with our earlier finding that agency connections promote larger temporary flows for this segment of the workforce.

I then use back-of-the-envelope calculations to quantify the aggregate efficiency gains from inter-regional labor intermediation via recruitment agencies. The expansion of agencies can account for over 10% reduction in the aggregate dispersion in regional wages by 2016. Furthermore, the labor reallocation results from population census data allow me to calculate predicted labor flows across the entire city system in China with and without agencies. I find that the inter-city networks of agencies reallocate around 27 million workers across space, and this translates into a 3.15% gain in aggregate output attributable to agency expansion by 2010[5]. Moreover, with the reduced form results on agency profits, I calculate how such aggregate gains from reducing misallocations are distributed between intermediaries and the rest of the economy (i.e., firms and workers). Strikingly, the estimates suggest that 40% of the output gains are captured by agencies as profits. This likely reflects the limited competition among different agency groups, since for many of the city pairs, there is usually only one agency group that connects them.

Lastly, from a long-run perspective, this paper explores how labor market intermediaries may endogenously emerge as the economy transitions from a highly misallocated stage towards the steady state of regional convergence. I write down a simple neoclassical model of regional growth which features mobility barriers as in Gennaioli et al. (2014). The steady state is efficient in this set up, suggesting that market frictions alone may not be a handicap to long-run efficiency, but the speed of convergence may be affected by access to markets or intermediaries. This resonates well with and contributes to a classic theme in the literature of growth and development, e.g., Barro et al. (1995) and Banerjee and Moll (2010). It also allows me to concentrate on how endogenous agency dynamics shape the convergence path. The core question I then explore is how the efficiency gains from labor reallocation via agencies can be amplified by regional capital accumulation during economic transition.

Calibrating the model to regional factor stock and productivities in China, this paper finds that agencies will be active in the initial periods where regional wage disparities are the highest and hence

[5]With some linear projection exercises, I also find there are likely larger output gains in later years.

there are larger opportunities to arbitrage. While agencies bring low-skilled workers from poor to rich regions, there are also higher incentives for capital accumulation in rich regions, and lower incentives for low-skilled worker to permanently migrate on their own. Reallocation of low-skilled workers by agencies generates substantial surplus gains compared to the no agencies counterfactual. The largest gain is observed at the initial periods, but agencies then exit as wage gap quickly shrinks. The surplus gains however remain persistent. The main idea is labor reallocations by agencies, in spite of its temporary nature, can speed up capital accumulation. Hence aggregate surplus is always on a higher path than the no agencies counterfactual until it reaches the steady state.

The paper then examines how alternative scenarios of the agency sector shape the transition dynamics toward the steady state. Reducing competition among agencies in both labor-sending and labor-receiving regions, thereby increasing their monopsony power over workers and bargaining power vis-à-vis firms, slows the transition, mirroring the effects of the 2013 Amendment to the Labor Contract Law, which raised entry barriers for agencies. In other words, greater agency competition in both regions generally enhances efficiency, although through distinct channels. Increasing agencies competition in labor-sending regions weakens agencies' labor market power and directly strengthens incentives for worker reallocation. In labor-receiving regions, higher agencies competition reduces their bargaining power with firms, allowing entrepreneurs to capture a larger share of the surplus from intermediation. Higher firm profits then accelerate capital accumulation, leading to faster convergence to the steady state. These results highlight that the distributional consequences of market competition have important dynamic efficiency implications. Moreover, workers tend to benefit more from agencies competition in the sending regions. Finally, I show that facilitating the transition of temporary agency workers into the permanent local labor force in high-wage regions can generate substantial efficiency gains and promotes a more equitable distribution of aggregate surplus. As more agency workers enter the advanced sector, the Lewisian process of labor reallocation accelerates: regional wage gaps shrink, agencies' rents decline, and agency workers earn higher wages. Taken together, these results reveal how agencies dynamics interact with the transition path of regional convergence, as well as how alternative structures of the agency sector determine who benefits more from reducing misallocations.

**Related Literature** This paper is related to the following strands of literature: first, following the seminal contribution by Hsieh and Klenow (2009) and Restuccia and Rogerson (2008), recent papers in the macro-development literature on misallocations start to (i) utilize natural experiments to estimate misallocations (Sraer and Thesmar, 2023; Bau and Matray, 2023), and (ii) look at their welfare and distributional effects (Baqaee and Farhi, 2019; Davila and Schaab, 2024; Atkin et al., 2024; Ghatak and Mookherjee, 2024). Brandt et al. (2025) show that the reduction of firms' entry barriers, driven by the shrinking SOE sector, can explain regional convergence in China. On the labor market, a few papers look at how firm size distortions lead to output loss (Garicano et al., 2016) and how spatial frictions shape labor misallocations across firms *within* regions (Heise and Porzio, 2023). Also related is Banerjee and Moll (2010) and Moll (2014) who argue that capital accumulation by entrepreneurs themselves may overcome financial constraints over time. This paper contributes to this literature by demonstrating the potential for profit-driven organizations to emerge endogenously and arbitrage away the distortionary wedges. This also shapes the functional distribution of gains from reducing labor misallocations.

Secondly, this paper provides a unique angle to understand the macroeconomic role of recruitment agencies in facilitating inter-regional labor mobility. Among the current papers on labor recruitment or staffing agencies and productivity, the most related paper is Bilal and Lhuillier (2025) who argue

that labor outsourcing in France leads to a 1.8% increase of aggregate output but 3.1% reduction in earnings and welfare of the low skill service workers. In another related paper, Bertrand et al. (2024) show that the increasing adoption of contract labor reduces the marginal cost of labor faced by large plants because of greater unionization, therefore reduces misallocation and can account for 7.6% of the manufacturing TFP growth in India. Estefan et al. (2024) find that domestic outsourcing of core manufacturing workers leads to lower labor compensation in Mexico, and the recent ban on such outsourcing practices increases labor income share, while reducing capital investment and increasing small firm exits. Felix and Wong (2024) finds that the 1993 outsourcing legalization in Brazil has lead to young workers replacing old workers in the security guards sector. Also related is Munoz (2024) who shows that EU's job posting policy increased temporary mobility of manual workers from low to high wage countries, with notably 37% profit increase in posting firms and 14% rise in wages following one posting event. This paper features the endogenous expansion of agencies, as well as imperfect competition of the agency sector, and discusses what this implies for regional wage convergence and aggregate gains via factor reallocations *across* regions.

The fact that agencies promote inter-regional labor flow also relates to the literature on internal migration in developing countries (see, for example, Young (2013), Munshi and Rosenzweig (2016), Bryan and Morten (2019), Lagakos et al. (2023)). I focus on temporary migration across cities in China and, in contrast with the spatial sorting mechanism emphasized in this literature, the migrant population in this paper is largely low-skilled[6]. More importantly, I use the staggered rollout of agencies as a natural experiment to identify their effects on labor mobility and provide new evidence on its aggregate implications, an area still under-explored in the migration literature (Lagakos (2020)).

Lastly, this paper also resonates with the earlier theories of economic development (Lewis, 1954; Banerjee and Newman, 1998), and growth theories of conditional convergence starting from Solow (1956). This paper contributes to it by showing how labor agencies serve as a bridge between the modern and traditional sector and speed up regional convergence of wages; and it highlights the role of competition among market intermediaries in shaping such dynamics.

The remainder of this paper is organized as follows: Section 2 describes the institutional background of labor agencies in China and presents a few additional descriptive facts constructed from different sources of datasets. Section 3 shows the core empirical results on the expansion of labor agencies, and its implications on factor reallocation and aggregate efficiency. In Section 4, I use a neoclassical growth model to illustrate how endogenous agency dynamics interacts with the transition path of regional convergence. Section 5 concludes.

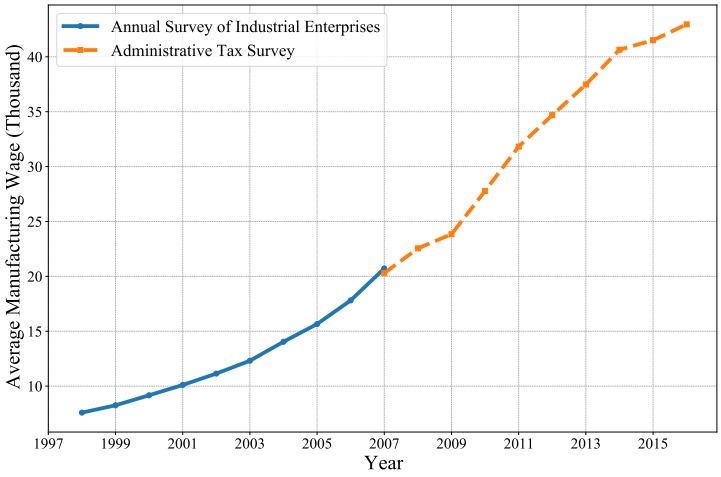
## 2 Background, Data, and Facts

The classic path of development as described in Lewis (1954) often entails rising wages, as the modern sector accumulate capital and absorb surplus labor from the traditional sector, a trajectory also observed in China. Figure 1 illustrates this process. Average manufacturing wages increased fivefold between 1998 and 2016, from under 10k to around 50k RMB. At the same time, the dispersions of wages across firms narrowed[7]: the coefficient of variation fell from 0.75 to 0.4, indicating wage convergence across producers. Interestingly, the sharp decline of wage dispersion alongside rising average wage in China's growth experiences contrasts with those of other developing economies such as Sub-Saharan Africa (Appendix Figure A.1), where wage growth is typically accompanied by greater

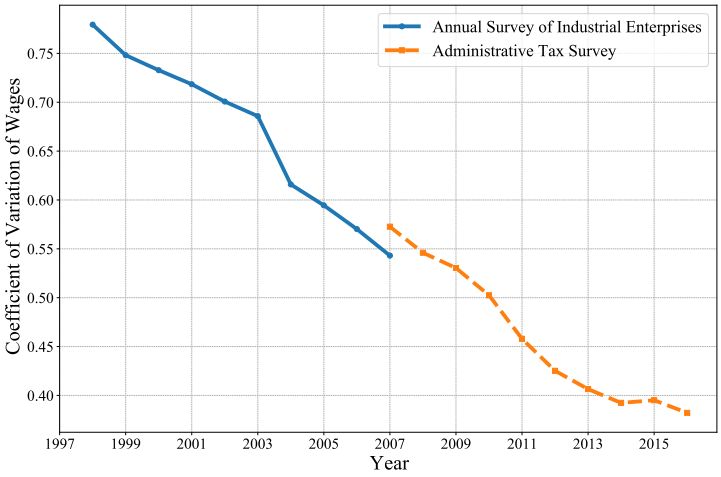
[6]See, for example, a recent paper by Becker and Hornung (2025) on how the coal-driven industrialization in Prussia raised wages for unskilled labor and lowered skill premium.

[7]Similar patterns of regional convergence in the manufacturing sector are also documented in Brandt et al. (2025).

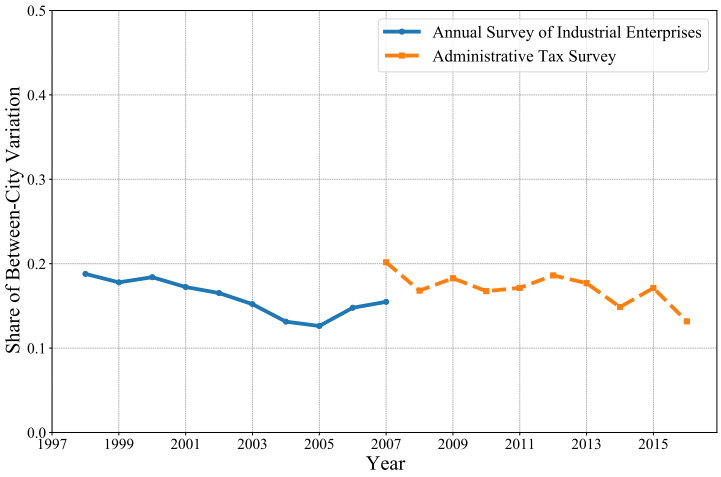
FIGURE 1: MANUFACTURING WAGES IN CHINA: 1998–2016




(A) AVERAGE WAGE




(B) WAGE DISPERSION




(C) CONTRIBUTION OF BETWEEN-CITY VARIATION

*Notes:* This figure shows average wages and the dispersion of wages in manufacturing sector in China over 1998–2016, based on the Annual Survey of Industrial Enterprises and Administrative Tax Survey. The top panel shows the annual average wage of manufacturing firms and the middle panel shows the coefficient of variation of firm wages. The bottom panel shows the share of between-city variation of firm wages in the national total variance, following the variance decomposition exercise:

$$\text{var}(w_{ic}) = \underbrace{\frac{1}{N}\sum_{c} N_c\left(\overline{w}_c - \overline{w}\right)^2}_{\text{between-city variance}} + \underbrace{\frac{1}{N}\sum_{c}\sum_{i\in c}\left(w_{ic} - \overline{w}_c\right)^2}_{\text{within-city variance}} .$$

where $w_{ic}$ is the average wage of firm $i$ in city $c$, $N$ is the total number of firms, $N_c$ is the total number of firms in city $c$, $\overline{w}_c$ is the average wage in city $c$, and $\overline{w}$ is the national average wage.

inequalities. In this paper, we will discuss how a specific labor market institution, namely, labor recruitment agencies, can contribute to China's experiences in declining wage dispersion.

A further decomposition of China's national variance of firm wages shows that the spatial wage disparities are persistent. Roughly 20% of the overall dispersion can be explained by between-city variations, and such share has remained stable. These patterns, i.e., rising wage levels, gradual wage convergence across firms yet persistent spatial wage disparities, consist of the empirical background from which labor recruitment agencies have emerged.

### 2.1 The Rise of Labor Agencies in China

Labor dispatch and outsourcing, where firms hire workers through recruitment and staffing agencies, has two distinct origins in China. In the late 1970s, foreign companies entering the Chinese market were required to hire workers through government-established agency platforms (Wang, 2019). Secondly, the large scale reform of state-owned enterprises (SOEs) in the late 1990s laid off millions of workers. Many of the SOEs therefore established labor agencies to facilitate the transition of these workers into new jobs. Later on, with the rising labor costs after 2000 (as shown above), domestic enterprises, both private and state-owned, began to adopt agency workers because of the flexibility and reduced costs in hiring and firing[8]. However, labor outsourcing practices were not legally recognized until the implementation of Labor Contract Law in 2008[9], which formalized the registrations of Labor Dispatch Agencies (LDA) in China.

According to the 2008 Labor Contract Law, agencies were required to have registered capital of at least 0.5 million RMB, and sign fixed term contracts with their workers, and provide monthly wages. In principle, they were allowed to send their workers to *temporary*, *auxiliary*, or *substitutable* positions. In practice, firms and agencies often exploited the ambiguity of these classifications, resulting in a much broader use of agency workers than the law originally intended[10](Liu, 2014). Firms sign contracts with agencies to obtain workers for auxiliary positions, which typically specify the number of workers and the total payment conditional on task completion (and we will explicitly model this in Section 4). This arrangement substantially lowers firing costs, since firms can simply return workers when necessary and request replacements from the agencies, an aspect commonly emphasized in the labor economics literature on staffing agencies (e.g. Autor (2003)). However, one important feature of China's labor outsourcing system is that agencies often send workers across cities, especially from surplus labor regions to more developed areas with higher demand[11]. Such inter-regional mobility constitutes the primary channel through which agency networks can reduce spatial labor misallocations, a mechanism that I emphasize in this paper. Notably, such temporary labor mobility for work purposes are not restricted by the *hukou* system, particularly for the recent decades that I study here.

[8]This is especially true for state-owned enterprises which are often subject to mandated quota for its formal employment. The use of agency workers enables them to bypass these restrictions.

[9]Therefore if a dispatched worker made a lawsuit case before 2008 of the company it worked for because of e.g, violation of contracts and delay of compensation payment, such cases might be identified as workers de facto directly working for the company as labor dispatch was not legally recognized by that time. In some ways, the introduction of Labor Contract Law helped to shield companies from such legal claims from the (dispatched) workers (Huang (2017)).

[10]The subsequent Labor Contract Law Amendment in 2013 and the No.22 Order of the Ministry of Human Resources and Social Security (MOHRSS) in China clarified these categories: " *temporary positions* refer to the positions that exist for not more than six months; *auxiliary positions* are non-main business that provide services for the positions of main businesses; and *substitutable positions* are those that can be held by substitute workers during a certain period of time when the employees of the employers who originally hold such positions are unable to work because they are undergoing full-time study, or on leave or for any other reason."

[11]This practice has sometimes triggered legal disputes due to regional differences in social benefits and compensation standards. In 2014, MOHRSS issued an additional regulation that agency workers' social benefits should be paid by either agency or firm according to local standards of the workplace.

The top panel of Appendix Figure A.2 plot the time series of monthly registrations of LDAs, based on the agencies dataset described below. The formalization of labor agencies in 2008 triggered a surge in registrations nationwide, with monthly entries peaking at around 750 in early 2013. It is also interesting to note that most of the post-2008 entrants are privately owned, which is the core distinction from the agencies established by the government of SOEs before 2008. Agencies are also perceived to employ a large fraction of the workforce: according to All-China Federation of Trade Unions (ACFTU), there are 37 million agency workers in 2011, which takes up 13.1% of the total workforce in enterprises in China.

The rapid rise of agencies also brought numerous labor disputes that expose firms' abusive use of agency workers, which include unequal pay for equal work, overtime work, limited career advancement channels and social insurance coverage[12] (Wang, 2019). These ultimately led to an amendment of the Labor Contract Law, which took effect on July 1, 2013. Focusing specially on labor agencies, it imposed several important restrictions: for examples, agencies were required to obtain an additional certificate from the Human Resources Department of the local government, and the minimum registered capital was raised to 2 million RMB. In addition, agencies were required to apply for separate certificates if they sought to establish new subsidiaries in different cities. The amendment also clarified the boundaries of "temporary, auxiliary, or substitutable" positions, narrowing the scope for ambiguity. Later in 2014, the MOHRSS imposed a further regulation that capped the share of agency workers at 10% of total employees for all firms. In practice, however, this has been hardly enforced: firms may circumvent the cap by signing "outsourcing" contracts, which is still not legally well defined, instead of dispatch contracts, even though workers are supplied by agencies in much the same way. Therefore it is hard to distinguish between labor outsourcing, labor dispatch, and temporary employment in the Chinese context. For this reason, my analysis focus on the broader spatial expansion of agencies and its aggregate impact on wages and efficiency.

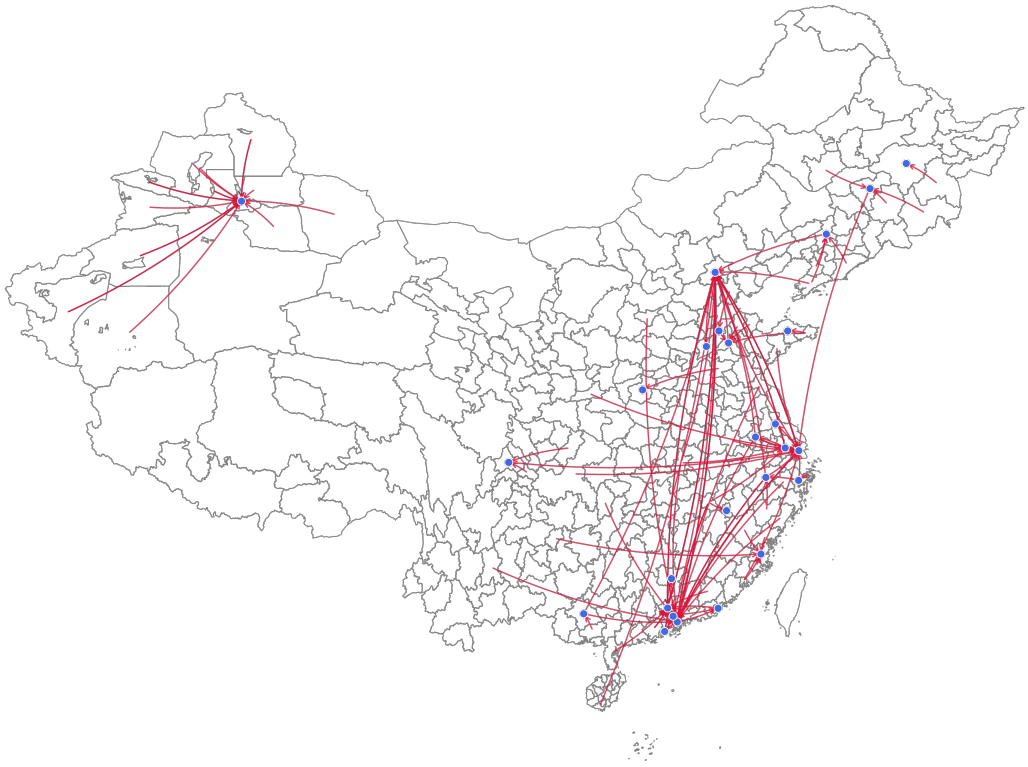
As in Appendix Figure A.2, the 2013 amendment resulted in a sharp drop of new LDA registrations, which then recovered again especially in late 2017 due to e.g., the rising demand for delivery workers from e-commerce platforms. Most of the post-amendment decline of new registrations stem from local agencies that are not connected to other cities via branches or headquarters; the inter-regional agency groups are largely unaffected. I plot the monthly registrations of such agencies (with branches or headquarters in other cities) in the middle panel of Appendix Figure A.2, which reaches around 200 in late 2018. Figure 2 illustrates the expansion of agency networks from 2005 to 2015: they began in relatively developed regions such as Beijing, Shanghai, and Guangdong before spreading to many other cities. Throughout the paper, I define a pair of cities being connected if any of the agency groups have active branches in both cities at the same time. Then in the bottom panel of Appendix Figure A.2, I plot the share of city pairs that are connected by agency groups, which rises from around 1% in 2007 to 17.5% in 2018, again suggesting the rapid expansion of agency networks across space. For the remainder of this paper, I will focus on the inter-regional agency groups that establish branches or subsidiaries (and promote labor reallocation) across multiple cities.

## 2.2 Data

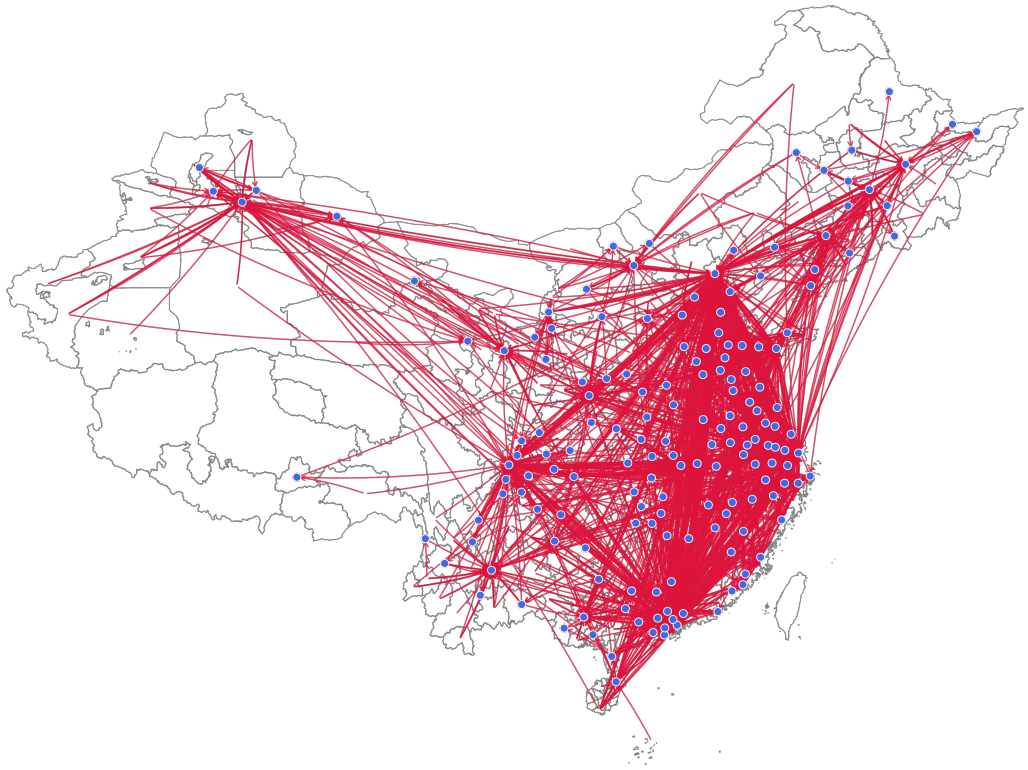
In this paper, I construct a novel comprehensive dataset that covers the near universe of LDAs in China that registered between 2000 and 2019. I obtain the agencies information on entry and exit,

[12] ACFTU issued a regulation in 2009 that required all LDAs to establish Trade Union and encourage agency workers to join. However, labor unions in China remain largely symbolic institutions, as they are largely politically controlled and often align with the employers in labor disputes (Tu (2017)).

FIGURE 2: AGENCY NETWORK



(A) 2005



(B) 2015

*Notes:* This figure shows expansion of agencies network from 2005 to 2015, based on my labor agencies dataset. The blue dots represent the headquarter cities for different agency groups, and the red lines connect headquarters with local branches.

location, ownership information, registered capital as well as business scope from one of the core business information platforms in China, *aiqicha*. Then I identify an enterprise as a LDA if its main business scope contains labor dispatch and or labor outsourcing, and their relevant characteristics of e.g., registered capital are consistent with the legal requirements mentioned above. Furthermore, I construct the spatial network of labor agency groups where the headquarter is identified as the LDA at the top of the ownership chain where the rests are local branches. I then manually classify the specializations of agencies by reading through their business scope descriptions.

Next, I combine the administrative tax survey dataset of firms, population census, as well as labor survey to study the impacts of agencies' spatial expansion. Following the current literature on worker migration in China (for example, Tombe and Zhu (2019), Fang and Huang (2022)), I use the 1% population *mini*-census in 2005 as well as the decennial census in 2010 to construct city level flow of workers (without changing initial *Hukou* registration location). I focus on the spatial reallocation of such *floating* population, which according to Chen et al. (2022), takes up 70% of all migration in 2000. The administrative tax survey of firms from 2007 to 2016 by the State Taxation Administration contains a repeated cross section of firms, covering various sectors[13]. I use this to calculate average wages and manufacturing marginal product of labor[14] for all cities. I also identify a sub-sample of labor agencies in the tax survey dataset which provides richer information on profits and employment. Lastly, I use China Labor Dynamics Survey (CLDS) in 2012, 2014, 2016, and 2018 to study how agencies expansion affect wages by educational background. The China City Statistical Yearbook provide information on average wages, fixed assets of the industrial sector, share of national GDP for each city, as well as total *hukou* population for my calibration and counterfactual analysis. It also contains data on number of high schools and their student-teacher ratios for each city, which are used to control for the local availability of educational resources. Lastly, I use road network data in 2000, 2010, 2014-2018 to calculate travel distance for each given city pair in China to capture the extent of transportation infrastructure development over the years[15]. This will be used as control variable in my difference-in-difference design on the consequences of labor agencies expansion. In Appendix Table A.1, I provide the summary statistics for the main variables I use for this paper.

### 2.3 Descriptive Facts

Even though they are perceived to employ a large fraction of the workforce, our knowledge of labor agencies in China remains scarce. Drawing on the above mentioned dataset, I am able to construct a few stylized facts consistent with the existing findings from similar contexts: for example, labor agencies in China pay on average lower wages than comparable business service firms, which is consistent with the general notion that agency jobs are often seen as "bad" jobs (Kalleberg (2011), Hatton (2011)). The dispersion of average wage among agencies is substantially higher than that in the manufacturing sector, with the left tail of the agency workers often identifying themselves as "second-class" workers. Apart from these, I now highlight the following facts that contribute new insights into how the agencies sector operates and motivate our further empirical investigation.

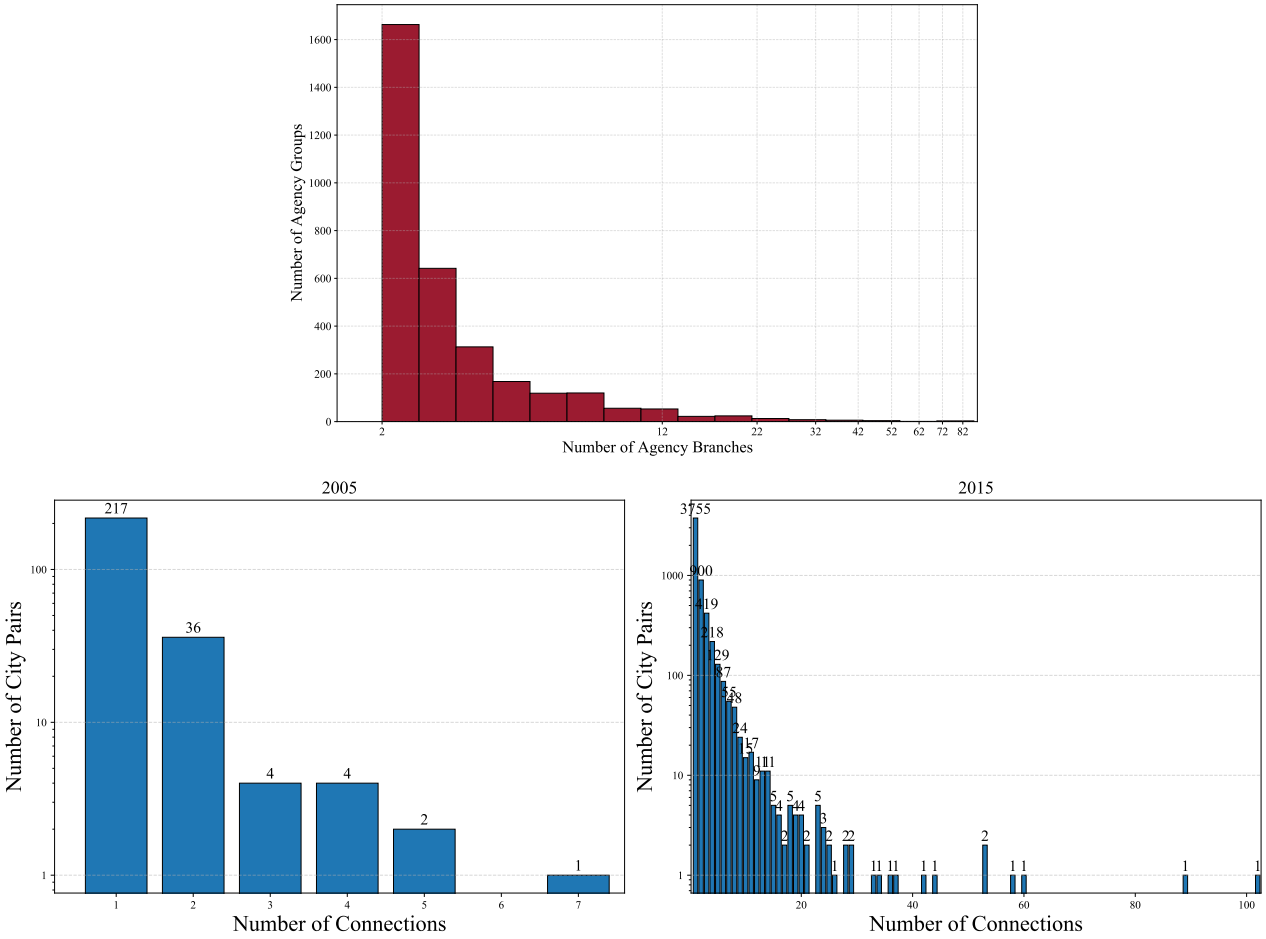
**Fact 1:** Agencies mainly covers business management, security, property management (including cleaning and security), as well as construction. Agency workers are less likely to have a college

[13] I also checked the consistency of the tax survey dataset with the more frequently used dataset of Annual Survey of Industrial Enterprises (ASIE) and the correlation of core variables (such as output and wages) between the two datasets is around 0.8-0.9, suggesting that the former dataset is reasonably reliable.

[14] I use the Cobb-Douglas production function as in Hsieh and Klenow (2009).

[15] For years where the road network data is missing, I use linear interpolation to calculate travel distance for city pairs.

Figure 3: Distribution of Agency Groups




*Notes:* The top panel shows the distribution of size of agency groups (i.e., the number of branches within an agency group). The bottom panel shows the frequency of city pairs that are connected (i.e., agency group have branches in both cities) by the corresponding number of agency groups in 2005 (left) and 2015 (right).

degree.

The bottom panel of Appendix Figure A.3 classifies agencies by task specialization. Around 75% of the labor agencies operate as general purpose human resources firms, and 10% are business management firms, which may encompass a wide range of services, ranging from (seasonal) assembly line labor in manufacturing firms, to cleaning, security, accounting, and translation. Besides, I also find more specialized labor agencies: 6.5% of the labor agencies focus exclusively on security services, and 3.9% of them are providing property management services (which include security and cleaning). These patterns suggest that agency workers are more likely to be concentrated in relatively low-skill occupations. I use the CLDS dataset to compare the characteristics of agency workers with those of directly hired workers, although the sample size of the former is fairly small due to missing observations in worker contract information. Nevertheless, only 13.2% of the surveyed agency workers obtained a college or above degree, which is significantly lower than the 21.3% of college graduates among direct hires. The fact that most of the agency workers are below college and working in low skill occupations foreshadows my later findings that the effect of agencies expansion on factor reallocation and wage convergence is stronger for below-college workers.

**Fact 2:** The agencies sector is characterized by a few large agency groups alongside many small ones. The largest agency group has over 80 branches across multiple cities in China, but the

majority of agency groups maintain only two branches. Many of city pairs are connected via only one agency group's spatial network.

In the top panel of Figure 3, I plot the distribution of number of branches among agency groups: the majority of agency groups only have two branches, and the largest agency group has over 80 branches spanning a number of cities in China. Then in the bottom panel, I plot the frequencies of city pairs that are connected (i.e., agency group have branches in both cities) by each given number of agency groups in respectively 2005 and 2015. In 2005, only 234 pairs of cities are connected by agency groups; and by 2015, around 10% of all pairs of cities (around 50k) are connected by the spatial branch network of labor agencies. However, around 65% of those connected city pairs are connected by only one agency group, while there are 6 pairs of cities being connected by over 50 agency groups. This implies there is potential specialization among agencies in connecting different pairs of cities, and the substantial bargaining power these specialized agencies have when they bring workers across the cities. Furthermore, in the bottom panel of Appendix Figure A.3, I use the sub-sample of agencies in the tax survey dataset to show the size (i.e., employment by headquarter of agency groups) distribution of agencies. This is again characterized by many small agencies that employ 10-100 workers and a few big agencies with more than 10000 workers. These are all important fact for interpreting my later results on the extent to which agencies can capture profits from reallocating workers across cities by arbitraging the wage differences.

Now I present my core empirical results on the expansion of labor agencies and its implications on labor misallocations in China.

# 3 Empirical Evidence

This section will present empirical evidence on (i) where labor agencies expand into given the pre-existing regional wage disparities, and (ii) how the spatial expansion of labor agencies affects inter-regional wage convergence and labor misallocations. Then we will use back-of-the-envelope calculations to translate the labor reallocation results into estimates of aggregate output gains.

## 3.1 The Expansion of Labor Agencies

Our first task is to explore where agencies choose to establish branches to recruit (or place) workers, following the 2008 Labor Contract Law which formally recognized their legal status. Intuitively, agency groups should start by exploiting the largest arbitrage opportunities by connecting cities that exhibit larger ex-ante wage disparities, i.e., those at opposite ends of the ex ante regional wage distribution. In this way they can promote labor reallocation from low wage cities with abundant surplus labor to high wage cities with strong demand for low-skilled or auxiliary workers. Throughout the paper, I use *ex ante* to denote time periods before the 2008 Labor Contract Law which formally recognized labor agencies[16].

To explore the relationship between ex ante wage disparities and the expansion decision of agency networks, I estimate the following specification at the city pair level:

$$connected_{mn} = \beta_1 \left|\Delta Wage^{pre}_{mn}\right| + \beta_2 distance_{mn} + \varepsilon_{mn}, \tag{1}$$

[16]This means 2007 in most cases, except that for the results on labor flow in Section 3.2.1, I will use city average wage in 2000 as *ex ante* wage because the census data covers 2000, 2005, and 2010. In fact, we can get similar estimation results if we use only the 2005 and 2010 census data and include the city average wage in 2005 as our measure of *ex ante* wage.

Table 1: ex ante Wage Disparity and Agency Network

| | Connected | | | Connection | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $\lvert\Delta Wage^{pre}\rvert$ | 0.201*** | 1.535*** | 0.830*** | 0.526*** | 1.487*** |
| | (0.00911) | (0.0627) | (0.0350) | (0.0324) | (0.0721) |
| Distance | -0.0872*** | -0.662*** | -0.364*** | -0.310*** | -0.801*** |
| | (0.00289) | (0.0202) | (0.0113) | (0.0152) | (0.0269) |
| Observations | 39930 | 39930 | 39930 | 39930 | 39930 |
| Methods | OLS | Logit | Probit | OLS | Poisson |

*Notes:* This table shows the regression results for the relationship between ex ante wage disparity between cities and expansion decision of agency network, where $\lvert\Delta Wage^{pre}\rvert$ is the absolute value of ex ante (i.e., 2007) logged average wage from the Chinese City Statistical Yearbook for each pair of cities. This is based on the sample of city pairs $mn$ that were not connected by any agency groups before 2007. *Connected* is a binary variable which equals 1 if both cities have ever been branched by any agency group during 2008–2018. *Connections* measures the number of agency groups between city pairs. Columns (1)–(3) use linear probability, logit, and probit models, respectively. Columns (4)–(5) estimate intensive-margin effects using OLS and Poisson models. Robust standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

where $connected_{mn}$ is a binary variable which equals 1 if there exists an agency group that has active branches in both cities $m$ and $n$ during 2008 - 2018, and $\lvert\Delta Wage^{pre}_{mn}\rvert$ is the absolute value of the ex ante difference of (log) average wage between the two cities in 2007. Moreover, I will also replace the outcome variable $connected_{mn}$ by $connection_{mn}$ which captures the number of agency groups that has active branches in the city pair $mn$, which captures the intensive margin of agency presence. $distance_{mn}$ is the (log) distance between cities $m$ and $n$. We expect $\beta_1 > 0$ because larger wage disparities across cities should incentivize agencies to establish branches to connect these cities.

Column 1 in Table 1 confirms this idea, which says 1% increase in ex ante wage disparity is associated with around 0.2 percentage point increase in the likelihood that the two cities are connected via any agency network in this linear probability model. It also says that a longer distance between the two cities discourage agency network to cover them. In column 2 - 3, I report the estimates of $\beta_1$ and $\beta_2$ respectively using logit and probit models. These also provide similar results that a higher ex ante wage disparity is associated with higher chance of the two cities being connected by an agency group. We can in principle specify $connected_{mn}$ by agency groups, and include group fixed effects into the specification: I have also studied the top 50 agency groups' expansion decision and find similar results[17].

I find higher ex-ante regional wage gap also incentivizes more agency groups to enter. In column 4 and 5, I respectively use OLS and Poisson models to estimate the relationship between $\lvert\Delta Wage^{pre}\rvert$ and $connection_{mn}$. Here, according to the later model, A 1% increase in the ex ante wage gap is associated with a 1.49% increase in the number of agency groups connecting the city pair. Taken together these results suggest that labor agencies establish their networks to connect cities at the both extremes of the ex ante wage distribution, and arbitrage the wage disparities caused by spatial barriers across local labor markets.

Another way to explore agency groups' location choices is to study the potential U-shaped relationship between branch location and ex ante city wages, i.e., agencies should be more likely to enter cities with the lowest and highest wages, but less so for those cities in the middle of the ex ante wage distribution. I therefore run the following specification for $branched_{rg}$ on a set of dummy variables

[17]This is because we have identified altogether around 4000 agency groups, many of these groups only establish branches in two or three cities, reflecting the limited informational or organizational capacity for these agencies to expand. Also including all agency groups will explode the size of the dataset.

of the decile group that city $r$ belongs to in the ex ante (i.e., 2007) distribution of city average wage from the Chinese City Statistical Yearbook:

$$branched_{rg} = \sum_{i=1}^{i=10} \beta_i \mathbb{1}\{r \text{ in decile } i\} + distance_{rg} + \delta_g + \mathbf{X_r} + \varepsilon_{rg}, \tag{2}$$

where $branched_{rg}$ is a binary variable which equals 1 if an agency group $g$ has ever established a branch in city $r$ during the ten years (i.e., 2008 - 2018) following the formalization of labor agencies. $\delta_g$ is the agency group fixed effects. $distance_{rg}$ is the (log) distance between the headquarter of group $g$ and city $r$. $\mathbf{X_r}$ is a set of *ex ante* city characteristics such as number of high schools, and student-teacher ratio in high school, as well as the unemployed population. These are included to deal with potential concerns that agencies might simply expand to places due to non-wage reasons such as limited quality and/or quality of educational resources, large unemployed workforce, etc. We will see that city labor wages remain an important determinant of agencies' branch location decisions, even after accounting for these non-wage city characteristics.

I report the coefficients estimates $\beta_i$ for all decile groups of equation (2) in Figure 4a, which uses the linear probability model where I treat the fifth decile group as the baseline group[18]. This result visualizes the U-shaped relationship between ex ante wage and agency expansion decision, with the first and the tenth decile group again having higher probability of being branched by agencies than the rest of the decile groups. To illustrate the *gradual* expansion of agencies across space over time, in Figure 4b, I separately plot the predicted probability of agencies establishing branches for cities at different ex ante decile group during years 2008-2013 and 2014-2018[19]. Here we find that in the first few years (i.e., 2008-2013) immediately following agencies being legally recognized, they started establishing branches in cities at the bottom two decile groups and the top decile group. Then they expanded across the country over 2014-2018, but again, cities at the bottom decile group and the top three decile groups see the largest increase in the predicted probabilities of agencies entries.

I contrast the branching choice of privately owned agencies and state-owned agencies (which takes up less than 5% of my agency sample) in Figures 4c and 4d, and an interesting pattern emerges here. As before, I find a U-shaped relationship between ex ante city wages and branching decisions for privately owned agencies; however, this pattern does not hold for state-owned agencies. Relative to cities in the fifth wage decile, these agencies exhibit a higher probability of entering the highest-wage cities, but no greater tendency to expand into lower-wage cities. These results corroborate the view that privately owned agencies are more likely than their state-owned counterparts to be driven by pecuniary motives in arbitraging inter-regional wage disparities.

In Appendix Table A.2, I report the coefficient estimates from a similar specification as equation (2) but only keeping dummy variables for the top and bottom 10% groups[20]. Column 1 is conditional whether the agency group had a branch in the city before 2008 (by including the variable Branched

[18] This is based on the sample of cities that had never been branched by any of the agency groups before 2007; therefore the results only reflect the *new* openings of agency branches after the 2008 Labor Contract Law.

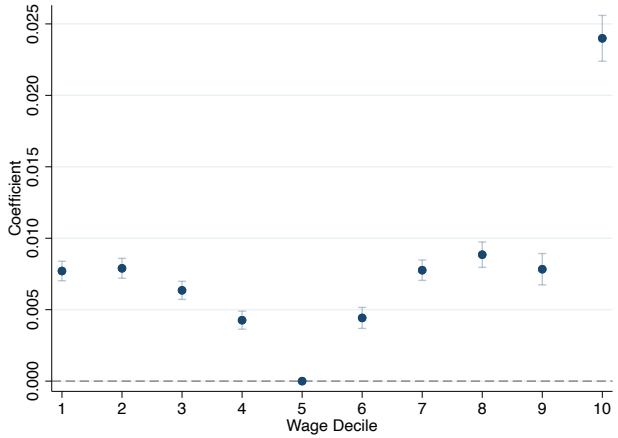
[19] Specifically, I run

$$branched_{rgt} = \sum_{i=1}^{i=10} \beta_{0i} \mathbb{1}\{r \text{ in decile } i\} + \sum_{i=1}^{i=10} \beta_{1i} \mathbb{1}\{r \text{ in decile } i\} \times \mathbb{1}\{t \text{ in } 2014-2018\} + distance_{rg} + \delta_g + \mathbf{X_r} + \varepsilon_{rgt},$$

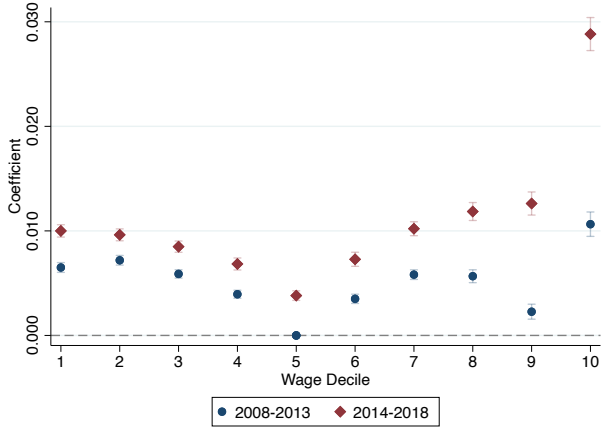
where $t$ is either 2008-2013 or 2014-2018. I plot the predicted probability for cities at different decile group being branched by agency groups separately for 2008-2013 and 2014-2018, namely, $\hat{\beta}_{0i}$ for the former and $\hat{\beta}_{0i} + \hat{\beta}_{1i}$ for the latter. Here I again treat the fifth decile group in 2008-2013 as the baseline group.

[20] Therefore the coefficients for these two decile groups capture the probability of a city being branched by agency groups *relative* to cities with ex ante wages between the 10th and the 90th percentiles.

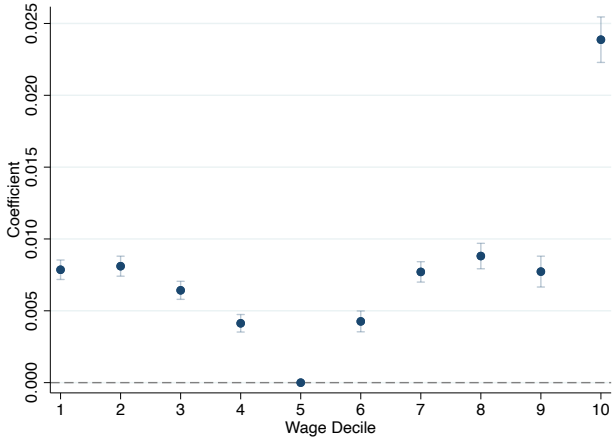
Figure 4: ex ante Wage Decile and Agency Expansion



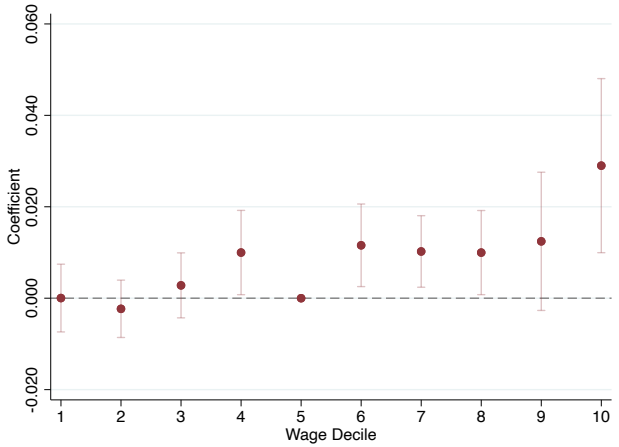

(a) Branched in 2008–2018




(b) Branched in 2008–2013 versus 2014–2018




(c) Branched by Private Agency




(d) Branched by State-owned Agency

*Notes:* Panels (a)–(c) report estimates from linear probability models of whether an agency group establishes a branch in a city, using decile groups of ex ante (i.e., 2007) city average wage (from the Chinese City Statistical Yearbook). All specifications control for distance between the city and the agency headquarters, the ex ante number of high schools in the city, and the ex ante student–teacher ratio in high school. Panel (b) reports estimates separately for 2008–2013 and 2014–2018. Panels (c) and (d) restrict the sample to privately owned agency groups and state-owned agency groups, respectively.

Before 2007). Cities in these two extreme decile groups of ex ante wage have a higher chance of being branched by an agency group. On the other hand, a larger geographical distance from agency headquarter city is associated with lower probability of being branched. Column 2 and 3 reports the coefficient estimates on the subsample such that no branch by the given agency group was established in the city before 2008, where I use a conditional logit model in column 3. Both columns report consistent results as in column 1 that agency groups are more likely to establish branches in cities at the two extremes of the ex ante wage distribution.

Our last task in this section is to demonstrate that such ex ante wage disparities can be translated into pecuniary benefits for agencies to capture. To explore this, I use the subsample of labor agencies from the administrative tax survey dataset and run average profits of agency headquarters on the maximum ex ante wage disparity among the cities that the agency group later expand into, and report the results in Table 2. In column 1, we find that on average 1% increase of the ex ante wage disparity is associated with 1.36% increase in agency profits. Column 2 includes the average size of the agency headquarters over 2008-2016 (when it is not missing from the tax survey dataset), which brings down the coefficient between ex ante wage disparity and agency profits to 0.75%. Interestingly, in column 3, I further include the number of agency groups that have expanded into the same city pairs with the maximum ex ante wage disparity, which is then negatively correlated with agency profits. This is consistent with the potential competition effect among agency groups in bringing down profits. Now 1% increase in ex ante wage disparity is correlated with 0.70% increase in agency profits. Also, a larger geographical distance between the cities tend to reduce agency profits. Column 4 further includes an interaction term between the number of (competing) agency groups and the ex ante wage gap. It shows that the competition effects are predominantly among city pairs with higher ex ante wage disparities, i.e., the marginal effect of number of competing groups on profits will reduce by 0.05% if there is 1% increase in the ex ante wage gap. In other words, the higher pre-existing wage disparities, the more (monetary) incentives for agency groups to expand and arbitrage; therefore there will be stronger competition among agency groups, which, as we will show below, will determine the distribution of efficiency gains from reducing misallocations.

### 3.2 The Impact of Agencies Expansion

The spatial expansion patterns reflect how ex ante wage disparities create pecuniary incentives for agencies. Building on this, we now examine how agencies expansion affects factor reallocation and wage convergence, and therefore contributes to efficiency improvement.

Our empirical strategy proceeds in several steps. We first document the relationship between agency connections and cross-city labor reallocation and wage convergence using two-way fixed effects regressions, and then use event-study estimators, including those proposed by Callaway and Sant'Anna (2021) and Arkhangelsky et al. (2021), to address the staggered rollout of agency networks and to assess pre-trends around the timing of connection formation. To link agency expansion more directly to misallocation reduction, we examine whether the effects of agency connections are stronger for city pairs with larger ex ante wage gaps, where the potential gains from reallocation are greater. Finally, we exploit the 2013 regulatory shock that imposed a higher minimum registered-capital requirement on labor agencies as an instrument for agency connections across city pairs. Together, these exercises provide complementary evidence that agency expansion causally facilitated labor reallocation from low-wage to high-wage cities and contributed to wage convergence.

We begin by presenting the core results on temporary labor flows across cities.

Table 2: ex ante Wage Disparity and Agency Profits

| | (1) Profits | (2) Profits | (3) Profits | (4) Profits |
|---|---|---|---|---|
| $\lvert\max \Delta Wage^{pre}\rvert$ | 1.356*** | 0.751** | 0.696** | 0.963*** |
| | (0.367) | (0.312) | (0.334) | (0.354) |
| Size | | 0.513*** | 0.504*** | 0.495*** |
| | | (0.0477) | (0.0476) | (0.0475) |
| Agency Groups | | | -0.0190*** | 0.00569 |
| | | | (0.00608) | (0.0139) |
| Distance | | | -0.0236 | -0.0176 |
| | | | (0.101) | (0.101) |
| Agency Groups× $\lvert\max \Delta Wage^{pre}\rvert$ | | | | -0.0501** |
| | | | | (0.0229) |
| Observations | 236 | 236 | 236 | 236 |

*Notes:* This table shows the regression results for the relationship between ex ante wage disparity of agency network and agency profits. Profits is the (log) average annual profits of headquarter of agency group over 2008-2016, based on the administrative tax survey dataset. $\lvert\max \Delta Wage^{pre}\rvert$ is the maximum ex ante (i.e., 2007) wage gap among cities that the agency group later expanded into during 2008-2018, where the city average wage is from the Chinese City Statistical Yearbook. Size is the (log) average total employment of headquarter of agency group over 2008-2016. Agency Groups is the number of agency groups that had established branches in the same cities that exhibit such maximum wage gap. Earliest is a binary variable which equals to 1 if the agency group is the earliest group that connect the two cities with $\lvert\max \Delta Wage^{pre}\rvert$. Robust standard errors reported in parentheses: *$p < 0.10$, **$p < 0.05$, *** $p < 0.01$.

### 3.2.1 Labor Reallocation

Utilizing the population census data in 2000, 2005 and 2010, I calculate $flow_{mnt}$, i.e., total headcount of individuals (in thousands) who had left their *hukou* registration location $m$ and stayed in city $n$ for more than one year (i.e., the "floating migrants" as discussed in the migration literature). I also calculate such total flow of by their education background (below or above college). Then combined with the agency network dataset, we can explore how agency expansion facilitated worker reallocation across regions (without changing their *hukou* registration) using the following specification:

$$Flow_{mnt} = \beta_1 Connected_{mnt} + \beta_2 Connected_{mnt} \times \Delta Wage^{pre}_{nm} + \lambda_{mn} + \gamma_t + travel_{mnt} + \varepsilon_{mnt}, \quad (3)$$

where $\Delta Wage^{pre}_{nm}$ is the ex ante (i.e., 2000) log wage difference between city $n$ and city $m$, i.e., $\Delta Wage^{pre}_{nm} = Wage^{pre}_{n} - Wage^{pre}_{m}$. $Connected_{mnt}$ is a binary variable which equals 1 if there exists an agency group that has active branches in both cities $m$ and $n$ in year $t$. As an extension, I will also include $Connection_{mnt}$, which is the number of such agency groups, as well as its interaction term with ex ante wage gap into the regression specification. This captures the intensive margins of agencies linkages across cities. We obtain the city average wage from the Chinese City Statistical Yearbook data. $\lambda_{mn}$ is the *directional* city pair fixed effects, which allows us to control for time in-variant factors specific to the directions of labor flow, and $\gamma_t$ is year fixed effects. I further include $travel_{mnt}$, i.e., road travel distance for the city pair calculated from road network data, to measure China's infrastructure development over the years and address potential concerns that this might be correlated with $Connected_{mnt}$ and also correlated with $Flow_{mnt}$. The estimate of $\beta_1$ therefore captures the non-wage reasons associated with labor flow, i.e., the network of agencies might also

Table 3: Agency Network and Labor Mobility Across Cities: 2000, 2005 and 2010

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Gross Flow | Gross Flow | Gross Flow | Gross Flow (College) | Gross Flow (Below College) |
| Connected | 8.439*** | 8.439*** | 5.193*** | 1.303*** | 3.726*** |
| | (0.994) | (0.966) | (1.268) | (0.312) | (0.986) |
| Connected$\times\Delta Wage^{pre}$ | | 29.34*** | 21.23*** | 4.105*** | 16.37*** |
| | | (3.113) | (4.049) | (0.888) | (3.211) |
| Connection | | | 2.110*** | 0.764*** | 1.294*** |
| | | | (0.580) | (0.142) | (0.462) |
| Connection$\times\Delta Wage^{pre}$ | | | 5.063*** | 0.801** | 4.167*** |
| | | | (1.494) | (0.314) | (1.247) |
| Travel | -5.404*** | -5.404*** | -5.156*** | -1.044*** | -3.919*** |
| | (1.749) | (1.725) | (1.741) | (0.364) | (1.349) |
| Observations | 200466 | 200466 | 200466 | 200466 | 200466 |
| *Fixed Effects* | | | | | |
| City Pair | Yes | Yes | Yes | Yes | Yes |
| Year | Yes | Yes | Yes | Yes | Yes |

**Notes:** This table reports the OLS estimates of the effect of expansion of agency network on worker mobility across cities. The results are obtained based on the sub-sample of population census in 2000, 2005 and 2010. $\Delta Wage^{pre}$ is the ex ante (i.e., 2000) log average wage difference between city $n$ and $m$. Gross flow from city $m$ to city $n$ (where $n \neq m$) is calculated as the total headcount (in thousands) for which the respondents reported to have lived in city $n$ (for more than one year) but their hukou registration is in city $m$. Gross flow of college (below college) labor is measured by such flow for which the respondents have a college (below college) degree. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. $travel_{mnt}$ is road travel distance for the city pair calculated from road network data. Standard errors reported in parentheses are clustered at city pair level: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

promote workers to work in different positions in Beijing or Shanghai with similar ex ante wages. We expect $\beta_2 > 0$ such that the expansion of agencies promote temporary labor flow from ex ante low wage to high wage cities.

Table 3 reports the regression results. The dependent variable is total gross flow for first three columns, and gross flow for college (below college) labor for the fourth (fifth) column. In column 1, I find that two cities being connected by an agency group is on average associated with around 8.4 thousands increase in the gross labor flow. Then in column 2, we see that such effect of labor flow increase will be stronger if there is a larger ex ante wage gap between the destination and origin city. Specifically, I find that one percent increase in the ex ante wage gap between cities $n$ and $m$ is associated with around 290 additional increase in the headcount of labor flow from $m$ to $n$ due to agency connection. I then include the intensive margins of agency connections in column 3, and now one percent increase in the ex ante wage gap is associated with around 210 additional increase in gross labor flow when cities are being connected by the agency groups. If such ex ante wage gap is zero, agency connection can also leads to on average five thousands increase in such labor flow, indicating that agencies might also facilitate temporary migration due to non-economic reasons. Furthermore, the intensive margins of agency connections, i.e., the number of agency groups connecting both cities, also contribute to increased labor flow, although with a smaller magnitude than the extensive margins. For example, conditional on the city pairs being connected, 1 percent increase in the ex ante wage gap is associated with 50 additional increase in gross labor flow following one additional agency groups connecting the city pairs.

Columns 4 and 5 report the estimation results separately for above and below college labor flow. We see that in general, there is a stronger reallocation effect among labor without a college degree: the higher is ex ante wage difference $\Delta Wage_{nm}^{pre}$, the expansion of agency network is found to generate a larger impact on the less-educated worker flows from city $m$ to city $n$. For example, 1% increase in the ex ante wage gap contributes to around 160 additional increase in below college worker flows from low wage to high wage city, contrasting with the 40 additional increase for above college labor. Correspondingly, we find again that the intensive margins of agency connections has a smaller impact on labor reallocation than the extensive margins, and such impact is also stronger among below college labor mobility. This confirms the primary role of labor agencies in organizing low-skilled surplus labor from ex ante low to high wage cities.

To summarize, these results provide a strong indication that the expansion of agency network has facilitated labor mobility, especially among those less-educated workers, from ex ante low wage cities to ex ante high wage cities. The next step is to show how the reallocations of factors can translate into changes in wage gaps.

### 3.2.2 Wage Convergence

If agencies reallocate labor from low wage to high wage cities, this will leads to wage convergence between the cities. To find empirical evidence for this, I first use the firm level tax survey dataset to calculate city average wage among manufacturing firms, and demonstrate if cities being connected by agency groups can see their average wage gaps reduced. In fact, one should also expect that such wage disparity reduction effects are more pronounced among city pairs that exhibit larger ex ante wage gap. In other words, since agencies have stronger incentives to reallocate workers among city pairs with larger wage gaps, there will be stronger effect of agency connections on narrowing wage disparities.

Table 4: Agency Network and Manufacturing Wage Gap: 2008 - 2016

| | (1) $\lvert\Delta Wage\rvert$ | (2) $\lvert\Delta Wage\rvert$ | (3) $\lvert\Delta Wage\rvert$ |
|---|---|---|---|
| Connected | -0.00900*** | 0.0503*** | 0.0418*** |
| | (0.00137) | (0.00173) | (0.00234) |
| Connected×$\lvert\Delta Wage^{pre}\rvert$ | | -0.274*** | -0.233*** |
| | | (0.00665) | (0.0126) |
| Connections | | | 0.00590*** |
| | | | (0.00107) |
| Connections×$\lvert\Delta Wage^{pre}\rvert$ | | | -0.0291*** |
| | | | (0.00765) |
| Travel | 0.129*** | 0.122*** | 0.122*** |
| | (0.00847) | (0.00811) | (0.00798) |
| Observations | 171250 | 171250 | 171250 |
| *Fixed Effects* | | | |
| City Pair | Yes | Yes | Yes |
| Year | Yes | Yes | Yes |

**Notes:** This table reports the OLS estimates of the effect of LDA expansion on (manufacturing) wage disparities across cities during 2008-2016. The results are obtained based on the administrative dataset of tax survey of firms. $\lvert\Delta Wage\rvert$ is the absolute value of difference of (log) average wage between two cities. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. Connection is the number of such agency groups. $\lvert\Delta Wage^{pre}\rvert$ is the absolute value of ex ante (i.e., 2007) log average wage difference. $travel_{mnt}$ is road travel distance for the city pair calculated from road network data. Standard errors reported in parentheses are clustered at city pair level: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

**Manufacturing Wages** The following two-way fixed effects specification allows us to explore this:

$$\lvert\Delta Wage_{mnt}\rvert = \beta_1 Connected_{mnt} + \beta_2 Connected_{mnt} \times \lvert\Delta Wage^{pre}_{mn}\rvert + \lambda_{mn} + \gamma_t + + travel_{mnt} + \varepsilon_{mnt}, \quad (4)$$

where $\lvert\Delta Wage_{mnt}\rvert$ is the absolute value of the difference of (log) manufacturing average wages between city $m$ and city $n$ at year $t$. $Connected_{mnt}$ is defined as before, and $\lambda_{mn}$ and $\gamma_t$ are respectively city pair and year fixed effects. For wage convergence, we should expect $\beta_2 < 0$ such that the wage gap reduction effects are stronger among city pairs that exhibit larger wage gap before the formalization of labor agencies in 2008. I report the results in Table 4. Note that I use the ex ante manufacturing average wage here and the dataset does not allow me to distinguish wages by agency and permanent workers. However, according to the Chinese accounting standards, agency workers and their compensations will also be included in their employer's accounting statement. Therefore the average wage among firms within the same city can be a reasonable representation of the workforce's average compensation.

Column 1 includes $Connected_{mnt}$ only and find that cities being connected by an agency group is associated with on average 0.94% reduction of their wage gap. I then include the interaction term in column 2, which supports the notion that the wage disparity reduction effects are stronger if the ex ante wage gap is larger: one percent increase in $\lvert\Delta Wage^{pre}_{mn}\rvert$ is associated with 0.28% additional reduction of wage disparity when there exists an agency group which include both cities into its spatial network[21]. In column 3, I further include the number of agency groups that have branches in

[21]The coefficient $\beta_1$ in specification (4) captures the (hypothetical) effect of agency groups connecting the city pair on its wage gap for those with the same ex ante wages. In data, the median of ex ante wage gaps is around 0.18, which then translates into an overall zero effect of agency expansion for such city pairs. Furthermore, I restrict the sample to

both cities, i.e., $Connection_{mnt}$, as well as its interaction with pre-existing wage gap, and find such intensive margin of agency connections are quantitatively less important than the extensive margin of $Connected_{mnt}$ in facilitating wage convergence. Here, after being connected by agency groups, 1% increase in the ex ante wage gap is only associated with around 0.03% marginal reduction of the regional wage disparity when an additional agency group comes in and connects the city pairs. This is intuitive because arbitraging the wage differences requires only one agency linkage, and additional agency linkages are less important for the law of one price to hold. Where such intensive margin of agency connection does matter, as we mentioned earlier, is the profits that agencies can appropriate when arbitraging such wage disparities.

**Event Study Specification** In parallel with the previous two-way fixed effects regressions, I now estimate the event study specification:

$$|\Delta Wage_{mnt}| = \sum_{h=-\tau}^{T} \beta_h \mathbb{1}\left[D_{mnt} = h\right] + \lambda_{mn} + \gamma_t + travel_{mnt} + \varepsilon_{mnt}, \tag{5}$$

where $\mathbb{1}\left[D_{mnt} = h\right]$ is the lead/lad indicator for the $h$-th year before/after there exists an agency network that connect the pair of cities $m$ and $n$, where $D_{mnt} = t - A_{mn}$ ($A_{mn}$ is the time period when the city pair $mn$ was first connected). Furthermore, to explore the regional convergence effects, I run this specification separately for city pairs that have ex ante high and low wage disparities[22], and we expect that the wage gap reduction effects should be more pronounced among city pairs with the former group. Furthermore, given our setting of staggered adoption and potential heterogeneity in treatment effects (and the potential negative weights problem), I estimate the above specification using the estimator proposed by Callaway and Sant'Anna (2021), and I will consider alternative staggered DiD estimators in Section 3.2.5.

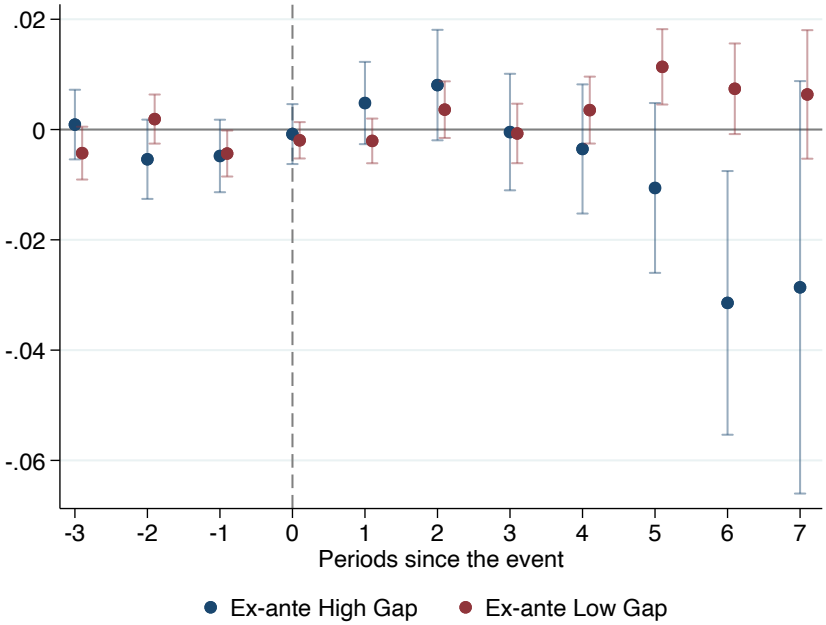
Besides, we have seen before that agencies selectively choose cities to establish branches and form their spatial networks. Estimating the above specification separately for city pairs with similar ex ante wage gap allows us to be more explicit on the causal interpretation of the wage convergence results. Namely, this implies a *conditional parallel trend assumption* such that for treated and control city pairs that exhibit similar ex ante wage disparity, they should evolve similarly over time.

I report the estimates of $\beta_h$ in Figure 5, which shows that the wage disparities among pairs of cities with ex ante high wage disparity start to decline when cities are connected via agency network, and there are minimal changes for the low wage disparity counterparts. I find an around 3% reduction after the sixth year when agency connection is established between the cities with ex ante large wage disparities. One can also see that the wage gap reduction effects become stronger in later years. Intuitively, agencies continue to promote inter-regional labor mobility over time, wage gaps keep narrowing and the connected cities get on the path of wage convergence, a core issue that I will return to in the model section. In contrast, among the city pairs with ex ante low wage disparity, there is no significant reduction of wage disparities. Furthermore, the pre-trend coefficients are close to zero and are all statistically insignificant, which increases our confidence that the conditional parallel trend assumption is likely to hold among city pairs that exhibit similar ex ante wage disparities.

city pairs with below median ex ante wage gap, and I find that the expansion of agencies has minimal impact over them. On average city pairs being connected by agencies is associated with 0.27% subsequent reduction of their wage gaps.

[22] Here I show the results for city pairs with ex ante wage gaps that are respectively above and below the 75th percentile of the wage gap distribution. One can obtain similar results by looking at the above and below median ex ante wage gap.

Figure 5: Event Study: Agency Network and Regional Wage Gap




*Notes:* The figure shows estimates from the event study specification of difference of (log) average (manufacturing) wage for city pairs on Connected (i.e., if there exist an agency group that has active branches in both cities), respectively for city pairs with above and below 75th percentile ex ante wage disparity, based on the methodology of Callaway and Sant'Anna (2021). I include road distance for each city pair as control. I also show the 95% confidence interval calculated using standard error clustered at city pair level.

**Comparable Sectors** Apart from the manufacturing sector, there are similar patterns of wage convergence in other comparable sectors as well. Appendix Table A.3 reports the estimates of specification (4) but the outcome variables are now the average wage disparities for comparable sectors that are also likely to heavily rely on agency workers, namely construction, hospitality, and logistics/warehousing. This is still based on the administrative tax survey dataset.

We can see that the expansion of agency networks also bring regional wage convergence for these sectors as well: for example, 1% increase in the ex ante wage difference is associated with respectively 0.08%, 0.11%, and 0.11% additional reduction in the sector specific wage gaps when agency networks connect the pairs of cities together. Furthermore, consistent with our previous results on the manufacturing sector, the intensive margins of agency connections also play a quantitatively less important role in bring down the wage disparities. Taken together these results underscore the generality of the wage convergence patterns brought by the expansion of agency networks.

**Wage by Education Background** So far I have shown that the expansion of agency network has facilitated regional convergence of *average* wage by sectors. But this may obscure important heterogeneities, and the most relevant to our analysis is how such convergence patterns differ by workers' human capital and/or skill levels. I utilize the CLDS dataset in 2012, 2014, 2016, and 2018 to calculate average annual wage for workers with or without a college degree for each city, as well as the corresponding wage gaps for city pairs. This allows me to run similar specifications as in (4) but now for skill specific wage[23].

The results are summarized in Table 5. Similar to the findings before, we can find strong wage convergence results for workers without a college degree: one percentage increase in ex ante wage gap is associated with 0.22% additional reduction upon up agency connection. In contrast, there are

[23]Note that the earliest CLDS survey data one could obtain is 2012, i.e., I cannot compute the ex ante wage gap $|\Delta Wage_e^{pre}|$ separately for college and below college workers. Instead, I calculate the average wage gap of city pairs from Chinese City Statistical Yearbook and use it as a proxy measure.

Table 5: Agency Network and Regional Wage Gap (by Education)

| | (1) $\lvert\Delta Wage_0\rvert$ | (2) $\lvert\Delta Wage_0\rvert$ | (3) $\lvert\Delta Wage_1\rvert$ | (4) $\lvert\Delta Wage_1\rvert$ |
|---|---|---|---|---|
| Connected | 0.0203 | 0.0189 | -0.00540 | 0.00106 |
| | (0.0172) | (0.0175) | (0.0320) | (0.0324) |
| Connected×$\lvert\Delta Wage^{pre}\rvert$ | -0.222*** | -0.210*** | -0.158** | -0.167** |
| | (0.0406) | (0.0428) | (0.0741) | (0.0760) |
| Connections | | 0.00000186 | | 0.00367 |
| | | (0.00251) | | (0.00256) |
| Connections×$\lvert\Delta Wage^{pre}\rvert$ | | -0.00924 | | 0.0117 |
| | | (0.00936) | | (0.0124) |
| Observations | 20119 | 20119 | 3895 | 3895 |
| *Fixed Effects* | | | | |
| City Pair | Yes | Yes | Yes | Yes |
| Year | Yes | Yes | Yes | Yes |

**Notes:** This table reports the OLS estimates of the effect of agency network expansion on wage (by education) disparities across cities in 2011, 2013, 2015, and 2017. The results are obtained based on the CLDS dataset (2012, 2014, 2016 and 2018). $\lvert\Delta Wage_e\rvert$ is the absolute value of difference of log average annual wage specific to educational background $e \in \{0, 1\}$ between two cities, where $e = 0$ stands for below college degree, and $e = 1$ is college degree. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. Connection is the number of such agency groups. $\lvert\Delta Wage^{pre}\rvert$ is the absolute value of ex ante (i.e., 2007) log average wage difference from the Chinese City Statistical Yearbook. Standard errors reported in parentheses are clustered at city pair level: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

no statistically significant changes to the below college wage gap for city pairs that have the same ex ante average wage. The magnitudes of the effects here are similar to what we found before in Table 4, despite the fact that Table 4 is based on different datasets and the manufacturing sector only. Furthermore, the intensive margins of agency connections have a quantitatively less important (in terms of magnitude) and statistically insignificant role in promoting regional convergence in low-skilled wages, again consistent with the previous results. However, when we move to column 3 and 4, I find smaller effects on wage convergence for the more educated workers with at least a college degree[24]: 1% increase in the ex ante wage gap is associated with around 0.16% additional reduction of the above college wage gap. These results suggest that wage convergence, as facilitated by the expanding labor agencies who were primarily organizing the low-skilled workers, might be *partial*, in the sense that labor wages at the lower spectrum of the skill distribution are converging, but less so for the more educated workers.

### 3.2.3 Factor Inputs

I have demonstrated that the expansion of agency networks leads to factor reallocation across regions, which in turn translates into convergence in factor prices (i.e., labor wages). The next results find similar convergence patterns for output per worker and capital-labor ratio among manufacturing firms across regions.

**Output per Worker** By calculating average output per worker (APL) across manufacturing firms for each city as well as the gaps of such average output per worker $\lvert\Delta APL\rvert$ between cities, I run the

[24]The drop in the number of observations is due to the lack of above college observations in the CLDS dataset. We obtain similar results if we restrict the sample to city pairs where both below and above college wage gap is none-missing.

same specification as in (4) and the results are summarized in Appendix Table A.4. As before, we find that the expansion of agency network leads to stronger effects of output per worker convergence among those city pairs that have larger ex ante APL gaps. 1% increase in the ex ante APL gap is associated with 0.16% additional reduction of APL gap across cities when they are connected by the agency groups. Furthermore, including the intensive margins of agency connections in column 3 again implies that this is quantitatively less important than the extensive margins to promote regional convergence of output per worker. Here, conditional on city pairs being connected, 1% increase in the ex ante APL gap is associated with only 0.02% additional reduction of the APL gap if one additional agency group connects the city pairs.

**Capital-Labor Ratio** We now examine how agency connections across cities affect regional gaps in manufacturing capital-labor ratio. Intuitively, consider the pair of an advanced (in Lewisian terms) city with high capital stock, low labor stock and high wage, as well as a backward city (i.e., the backward sector) with low capital stock, high labor stock and low wage. When agencies connect these cities, they bring labor from the latter to the former, which will automatically bring down the capital-labor ratio gap. However, this effect will be at least partially offset if the rich city further increase capital input as a response. Appendix Table A.5 presents the result. Here, we find one percent increase in the ex ante capital-labor ratio gap is associated with around 0.07% additional reduction of the regional capital-labor ratio gap when agency groups establish branches across cities. The magnitude is slightly lower than the convergence results for wages and APL. We do not find evidence that the intensive margin of agency connections contribute to further reduction of the capital-labor ratio gap.

### 3.2.4 IV Estimation

I can also instrument $Connected_{mnt}$ using the policy shock of the Amendment to Labor Contract Law in 2013. As discussed in Section 2.1, it imposes the minimum registered capital requirement, and requires such registered capital to be fully *paid in*. The administrative barriers for the agency sector also increase because agencies also need to apply for separate certificates from local government to open up new subsidiaries in different cities. This policy shock therefore provides important variation for the operation of agency branches across cities, depending on whether the branch is a *local office* or a *subsidiary*. By legal definition, local office is within the same legal entity with the agency headquarter/owner, so it does not need to register with additional registered capital. This is in contrast with subsidiary which now need to have fully paid-in capital when they register and also apply for additional certificate after the 2013 Amendment. Such regulatory differences between *local offices* and *subsidiaries* did not exist prior to 2013 and led to substantial exits and sharp declines in subsidiary registrations thereafter. Moreover, the choice of subsidiaries versus local office prior to 2013 by agency groups is largely driven by owners' risk and organizational preferences instead of regional factors, e.g., since local offices are within the same legal entity, their debts need to be repaid by the headquarters or agency owners.

Therefore, city pairs connected by agency *local office* as opposed to *subsidiaries* before 2013 are more likely to remain connected thereafter. To capture how exposed each city pair was to such negative policy shock, I calculate

$$\Theta_{mn} = \frac{\#\{\text{agency groups connecting } (m,n) \text{ through } \textit{local offices}\}}{\#\{\text{agency groups connecting } (m,n) \text{ in 2012}\}} \tag{6}$$

Table 6: First Stage: 2013 Amendment and Connectedness

| | $Connection_{mnt}$ |
|---|---|
| $\Theta_{mn} \times Post_t^{2013}$ | 2.115*** |
| | (0.082) |
| Observations | 604,406 |
| F-statistic | 671.28 |
| *Fixed Effects* | |
| City Pair | Yes |
| Year | Yes |

*Notes:* $\Theta_{mn}$ is the share of agency connections via local offices for city pair $mn$. $Post_t^{2013}$ is a binary variable which equals to one if year $t$ is after 2013. Standard errors reported in parentheses are clustered at city pair level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

and I can therefore instrument $Connection_{mnt}$ using $\Theta_{mn} \times Post_t^{2013}$, where $Post_t^{2013}$ is a binary variable which equals to one if year $t$ is after 2013. I report the first stage result in Table 6, which shows that the more city pairs are connected through local offices, the more connections remain in the city pair after 2013. Reassuringly, in Appendix Table A.6, I show that, after absorbing province-pair fixed effects, predetermined city-pair characteristics such as gaps in GDP per capita, agricultural employment share, and population, as well as road distance and pre-period labor flows, are not systematically correlated with $\Theta_{mn}$. These results strengthen the empirical case for the validity of the instrument by showing that $\Theta_{mn}$ is not systematically driven by pre-existing economic differences across city pairs.

We now report the results of the consequences of agencies expansion on labor flow and regional wage convergence in Table 7, where we instrument $Connection_{mnt}$ by $\Theta_{mn} \times Post_t^{2013}$. Panel A documents the effect of agency expansion on gross labor flow using population census data in 2000, 2005, 2010, and 2015. With the full sample in column 1, I find that city pairs being connected by agencies is on average associated with around 0.9 thousand increase in the gross labor flow from city $m$ to city $n$. Columns 2-4 are respectively based on the sample of city pairs where the destination city $n$ has a higher (lower) ex ante wage than origin city $m$, i.e., columns 2 and 3 (columns 4 and 5), and the absolute value of ex ante wage gap $|\Delta Wage_{mn}^{pre}|$ is above (below) the 75th percentile of the wage gap distribution, i.e. columns 2 and 5 (columns 3 and 4). A comparison across columns (2)–(4) indicates that agency expansion has a stronger effect on labor flows from low-wage to high-wage cities, especially among city pairs with larger ex ante wage differentials. For instance, for city pairs in the top quartile of the ex ante wage gap distribution, agency presence in the city pair is associated with an average increase of roughly 2 thousand workers moving from low-wage to high-wage cities. By contrast, labor flows from high-wage to low-wage cities decrease by approximately 0.16 thousand workers when the ex ante wage gap is large.

In Panel B of Table 7, I report the IV estimates of the impact of agencies expansion on manufacturing wage gap. Here, again consistent with what we have seen before, city pairs being connected by agencies is associated with on average 0.4% reduction of the manufacturing wage gap for city pairs that exhibit above 75th percentile ex-ante wage disparity. In contrast, city pairs with lower ex-ante wage gap witnessed minimal impact from agencies expansion.

Table 7: Agency Expansion, Labor Flow, and Manufacturing Wage Gap: IV Estimation

**Panel A: Gross Labor Flow**

| | (1) Full Sample | (2) Low→High High $\lvert\Delta Wage^{pre}_{mn}\rvert$ | (3) Low→High Low $\lvert\Delta Wage^{pre}_{mn}\rvert$ | (4) High→Low Low $\lvert\Delta Wage^{pre}_{mn}\rvert$ | (5) High→Low High $\lvert\Delta Wage^{pre}_{mn}\rvert$ |
|---|---|---|---|---|---|
| Connection | 0.964*** | 2.077*** | 1.635*** | 0.285*** | -0.164*** |
| | (0.160) | (0.654) | (0.351) | (0.0593) | (0.0618) |
| Travel | -2.872*** | -19.69** | -1.684** | -0.0393 | 2.071*** |
| | (0.958) | (7.739) | (0.806) | (0.225) | (0.627) |
| Observations | 267,288 | 33,412 | 100,208 | 100,208 | 33,412 |

**Panel B: Manufacturing Wage Gap**

| | (1) Full Sample | (2) High $\lvert\Delta Wage^{pre}_{mn}\rvert$ | (3) Low $\lvert\Delta Wage^{pre}_{mn}\rvert$ |
|---|---|---|---|
| Connection | -0.000385 | -0.00408* | 0.0000510 |
| | (0.00109) | (0.00217) | (0.00100) |
| Travel | 0.109*** | 0.171*** | 0.0989*** |
| | (0.00717) | (0.0195) | (0.00773) |
| Observations | 274,491 | 42,817 | 128,433 |

**Notes:** Panel A reports IV estimates where the dependent variable is gross bilateral labor flows between cities. Columns (2)–(5) split the sample by direction (low-wage to high-wage vs. high-wage to low-wage) and by the magnitude of the ex ante wage gap $\lvert\Delta Wage^{pre}_{mn}\rvert$. This is based on gross labor flow using population census data in 2000, 2005, 2010, and 2015. Panel B reports IV estimates where the dependent variable is the absolute value of manufacturing wage gap $\lvert\Delta Wage_{mnt}\rvert$. This is based on the tax survey dataset 2007-2016. High/low ex ante wage gap are defined by whether the city pair $mn$ is above and below the 75th percentile of $\lvert\Delta Wage^{pre}_{mn}\rvert$ among city pairs. City pair fixed effects and year fixed effects are included. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### 3.2.5 Further Results and Robustness

We now consider a few extensions of the empirical results, which shows (i) there are no pre-existing differential trends of worker flow for connected city pairs; (ii) how the empirical patterns of wage convergence show up in alternative DiD estimators and when considering regional spillovers; and (iii) how agency connection affects the temporary net flow of labor and changes in *Hukou* population across cities.

**Differential Trends in Worker Flow** One potential concern with respect to the labor reallocation results is that there may be pre-existing differential trends of worker flow for the connected city pairs (e.g., agencies may arrive simply because there are already larger worker flow between city pairs), which may bias the Difference-in-Differences results.

To address this concern, I use the population census in 2000 and calculate the gross labor flow across city pairs, which is then combined with the 2005 gross labor flow data. I then estimate

$$Flow_{mnt} = \beta_1 Treat_{mn} + \beta_2 \mathbb{1}[t = 2005] + \beta_3 Treat_{mn} \times \mathbb{1}[t = 2005] + \varepsilon_{mnt}, \tag{7}$$

where $Treat_{mn}$ is a binary variable which equals to one if city pair $mn$ is connected in 2010, and zero otherwise[25]. $\mathbb{1}[t = 2005]$ is an indicator variable which equals to one if year is 2005. It then follows that $\beta_3$ captures whether the treated city pairs have a differential trend (between 2000 and 2005) relative to the control city pairs.

I report the result in Appendix Table A.12. Here I could not find evidence that there exists pre-existing differential trends among treated city pairs between 2000 and 2005. The first column report the results for the entire sample of city pairs, and the next two columns are for city pairs with ex ate high and low (i.e., above and below the 75th percentile) wage gap. Our estimates of $\beta_3$ are not statistically different from zero, which increases our confidence that the labor reallocation results are not driven by pre-existing trends in labor flows before agencies arrive.

**Alternative Difference-in-Differences Estimators** In Appendix Figure A.6, I show the estimates of the event study specification of (5) but with the estimator proposed by de Chaisemartin and D'Haultfoeuille (2024). Here, among city pairs with ex ante low wage disparity, I find there is minimal impact on the subsequent wage gap when they are connected via agency network. The reduction of regional wage disparities is more pronounced among city pairs with above the 75th percentile ex ante wage gap, consistent with our existing results. Here, after the seventh year of being connected by agencies, city pairs with high ex ante gap observe 3% reduction of the wage disparity.

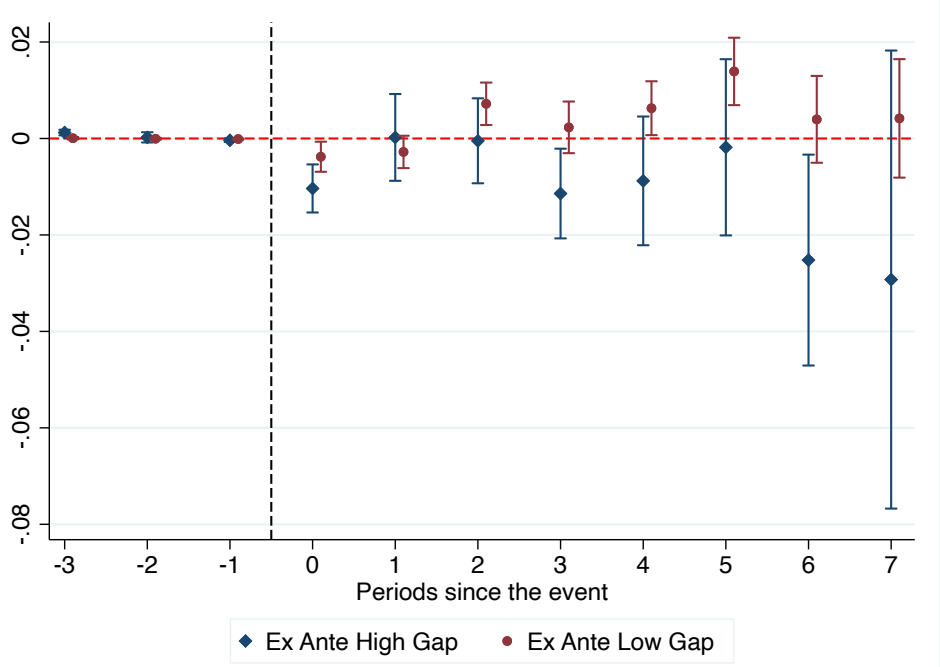
I also estimate the event study specification using the synthetic DiD estimator proposed by Arkhangelsky et al. (2021). The core idea of this method is to calculate unit and time specific weights and then estimate the difference-in-differences specification with weighted units and time periods such that the average pre-treatment outcomes of the treatment and control groups are approximately parallel (hence the confidence intervals of the pre-trend coefficients will be small by construction). Figure 6 summarizes the result, where the top (bottom) panel is for city pairs with high (low) ex ante wage gap.

Here, for city pairs with ex ante wage gap above the 75th percentile[26], we observe subsequent reduction of regional wage disparities after they are connected by agency branches, and such effect is stronger in later years: in the sixth or the seventh year after agency groups established the linkages, I

[25] I exclude the city pairs that are treated in 2000 and 2005 (around 400 observations)

[26] Again, I can find similar results for different cutoffs such as the median in the distribution of ex ante wage gap.

Figure 6: Agency Network and Regional Wage Gap: Synthetic DiD Estimates




*Notes:* This figure shows estimates of the event study specification of difference of (log) average (manufacturing) wage for city pairs on Connected (i.e., if there exist an agency group that has active branches in both cities), respectively for city pairs with ex ante above and below 75% wage disparity, based on the methodology of Arkhangelsky et al. (2021). I also show the 95% confidence interval calculated using bootstrapped standard error.

find around 3% reduction of the wage disparity, and the magnitude is similar to our previous staggered DiD estimates [27]. On the contrary, for city pairs with ex ante low wage gap, I could not find evidence for further reduction of their wage disparities. This result, based on weighted treatment and control city pairs having parallel pre-treatment outcomes, further consolidates the findings that the expansion of agencies across the country leads to reductions in regional wage disparities.

**Spillovers** We now discuss how potential sources of spillovers may affect our wage convergence results. We first consider a type of *geographical* spillovers such that when agencies establish branches in an ex ante high-wage city $n$ and a low-wage city $m$, workers from nearby low-wage city $k$ close to $m$ may see the opportunity and commute to $m$ for a job that send them to $n$. This implies that control city pair $(k, n)$ may be affected by the treatment of $(m, n)$, a violation of the SUTVA assumption that leads to under-estimation of the treatment effect. To address this concern, for each treatment city pairs $(m, n)$, I find all the neighboring cities $k$ of the low-wage counterpart $m$, $k \in N_m$, and I define the potential spillover control group as

$$S^G_{mnt} = \left\{(k, n) \mid k \in N_m, Wage^{pre}_k < Wage^{pre}_m, Connected_{knt} = 0\right\}.$$

This includes all control city pairs $(k, n)$ such that neighboring city $k$ of $m$ has a lower ex ante wage than city $m$: it is more plausible or attractive for workers in $k$ to commute to $m$ to seek an agency job.

Removing these potential spillover control groups, I re-estimate equation (4) and report the results in Appendix Table A.7. Here, there is a slightly larger effect of wage gap reduction compared to Table 4: one percent increase in the ex ante wage gap is now associated with 0.24-0.28 percent additional

[27] Note that however the synthetic DiD methodology requires the panel data to be perfectly balanced, but this is not required for the staggered DiD estimators proposed by Callaway and Sant'Anna (2021) and de Chaisemartin and D'Haultfoeuille (2024)

reduction in the regional wage disparities. I also run the event study specification and report the results in Appendix Figure A.4, where again we find slightly larger estimate of the wage gap reduction effects in the sixth year of agencies connecting the city pairs.

Such *geographical* spillovers may also affect our labor flow results. In Appendix Table A.8, I exclude the control city pairs specified in $S^G_{mnt}$ (for each treatment city pair $mn$) and present the estimation result for (3). Here we identify stronger effect of labor flow: for example, when any of the agency groups connect the city pairs, this is associated with on average 7 thousands worker flow (column 1). Then in the next two columns, 1 percent increase in the ex ante wage gap is now associated with around 100-200 additional worker flow. Similar to our findings before, the expansion of agencies networks has a stronger effect over below college flow than the above college flow. To summarize, these results confirms the idea that the type of geographical spillovers discussed above tend to bias downward the true treatment effect of agency expansion.

The second type of spillovers concerns *equilibrium price effects*. Namely, when agency connection brings down wage gap of city pair $(m, n)$ by raising the wage of $m$ and reducing the wage of $n$, then all city pairs $(k, m)$ and $(s, n)$ (where $k, s \neq n, m$) may (by definition) be affected. Since it is hard to find enough pure control city pairs with no agency presence, I define the following measures that account for how exposed/connected each city pair is to the expanding agency networks in the rest of the city system – this allows us to check how sensitive the results can be to the intensity of spillovers through price effects. For each city $m$ (and city pair $(m, n)$), I calculate a weighted sum of how many cities are connected to it:

$$\Psi^n_{mt} = \sum_{k \neq m,n} w_{mk} \times Connected_{mkt}, \tag{8}$$

where $w_k$ is the ratio of population of city $k$ over the distance between $k$ and $m$. Similar to measures of market access in the spatial economics literature (Donaldson and Hornbeck, 2016), I assume there is likely to be larger spillover effect if $m$ is connected via agencies to a larger city $k$ within shorter distance. I can then calculate how exposed each city pair is to the rest of the city system by taking $\hat{\Psi}_{mnt} = \Psi^n_{mt} + \Psi^m_{nt}$.

Then in Appendix Table A.9, I re-estimate equation (4) and report the results by removing the control city pairs that are highly exposed/connected to the rest of the city system (i.e., above the 90th percentile in the distribution of $\hat{\Psi}_{mnt}$). Here we find that one percent increase in the ex ante wage gap is now associated with around 0.29-0.33% additional reduction of the wage disparities. In the similar vein, I show the even study estimates in Appendix Figure A.5 using the rest of the sample after removing the highly exposed control city pairs. Again, there are minimal changes to the city pairs with ex ante low wage gap; but for the ex ante high wage gap city pairs, we now see slightly larger effects of wage disparity reduction than in Figure 5. For example, in the sixth year following agency groups connect such city pairs, we observe a 4% reduction of the wage disparities. This is again consistent with the idea that the equilibrium price effects specified above are likely to lead to under-estimation of the true treatment effect, especially for control city pairs with high ex ante wage disparity.

**Average Wage from City Statistical Yearbook** We can also use the city average wage data from the Chinese City Statistical Yearbook to estimate specification (4). We find similar patterns of wage convergence as before, which is summarized in Appendix Table A.10. Interestingly, now the wage gap reduction effect is smaller in magnitude compared to the manufacturing wage results, i.e., one percentage increase of the ex ante wage gap is now associated with 0.11-0.12% additional reduction

of the wage gap when agencies connect the city pairs. This is consistent with the idea that agencies only reallocate a particular fraction of the labor force across space, so one will observe smaller effect when looking at the overall average wage.

***Hukou* Population Change** The last result of this section concerns whether agency connection leads to changes in the distribution of *Hukou* population across cities. Although agency jobs do not directly confer *Hukou* status and agency workers therefore have limited access to local public resources, they may nevertheless obtain longer term residence in high-wage cities (and the *Hukou* status) through changes in employment contract, job transition, or marriage, etc. To explore this angle, I use the *Hukou* population data for each city in the Chinese City Statistical Yearbook, and calculate *Hukou* population gap for each city pair. I report the city pair level regression results in Table A.11, where one percent increase in the ex ante average wage gap is associated with 0.06-0.08 percentage additional increase in the *Hukou* population gap following agencies connect the city pairs[28]. This is consistent with the notion that agencies may promote temporary mobility of workers from low-wage to high-wage cities, and these workers may later on consolidate their residence by obtaining *Hukou* status.

## 3.3 Aggregate Gains

In this section, we use the empirical results on labor reallocation and wage convergence across cities and back-of-the-envelope calculations to quantify the aggregate efficiency gains from the expansion of labor agencies.

### 3.3.1 Reduction in Manufacturing Wage Dispersion

We start by calculating how the expansion of agencies across the country contribute to reduction in overall dispersion in regional manufacturing wages, one of the core metric of aggregate misallocations. Let $w_{i,t}$ denote the average (manufacturing) wage in city $i$ at year $t$. Then the national variance of wages can be expressed by the sum of squared pairwise wage difference:

$$\text{var}\left(w_{it}\right) = \frac{1}{N^2}\sum_{i<j}\left(w_{it} - w_{jt}\right)^2. \tag{9}$$

Then using the regression results on city average wage in the manufacturing sector[29], I calculate the predicted value of $|\Delta Wage_{mnt}|$ with agencies for each year from 2008 to 2016 by plugging in the corresponding $Connected_{mnt}$ and $Connection_{mnt}$, and that for the no agencies counterfactual by setting $Connected_{mnt} = 0$ and $Connection_{mnt} = 0$. I can then calculate the national variance of wages using (9) with and without agencies by summing up the squared fitted value of $|\Delta Wage_{mnt}|$.

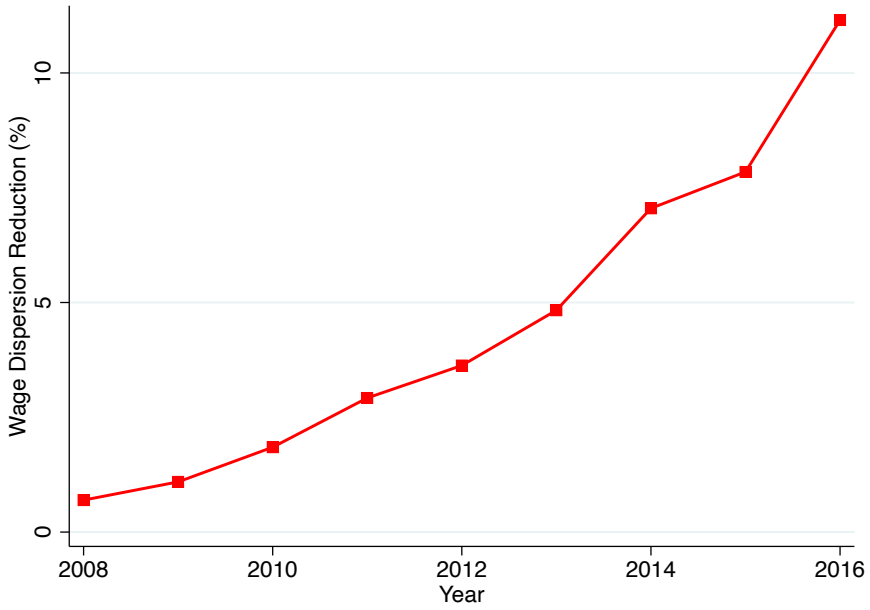
Figure 7 presents the reduction of wage dispersion due to agencies relative to the no agencies counterfactual. As agencies continue to expand across the country, by 2016, their connections across cities account for over 10% of the aggregate dispersion in regional wages, rising from around 2% in 2010.

[28] Note that this is a smaller magnitude compared to the temporary labor flow results based on the population census data.

[29] Specifically, I run

$$\begin{aligned}|\Delta Wage_{mnt}| \quad = \quad & \beta_1 Connected_{mnt} + \beta_2 Connected_{mnt} \times |\Delta Wage^{pre}_{nm}| \\ & + \beta_3 Connection_{mnt} + \beta_4 Connection_{mnt} \times |\Delta Wage^{pre}_{nm}| + \lambda_{mn} + \gamma_t + \varepsilon_{mnt}.\end{aligned}$$

FIGURE 7: REDUCTION OF WAGE DISPERSION




*Notes:* This figure shows the reduction of wage dispersion due to agencies, i.e., $\frac{\text{var}(w_{it}^c)-\text{var}(w_{it})}{\text{var}(w_{it}^c)}$, where $w_{it}^c$ denotes average manufacturing wage in the no agencies counterfactual. This is based on the tax survey dataset of manufacturing firms (2007-2016).

### 3.3.2 Output Gains

We can also quantify the output gains brought by the expansion of agencies. I take the standard constant returns to scale Cobb-Douglas function as the regional aggregate production function: (i) $Y = AK^{\alpha}L^{1-\alpha}$ where $L$ is the total labor in the region, and I also allow for (ii) $Y = AK^{\alpha}(L_U)^{\beta}(L_S)^{1-\alpha-\beta}$ where $L_U$ and $L_S$ are respectively low-skilled (below college) and skilled (college) labor. We also assume that regional capital stock and productivity are fixed at least for the short run. Then using $Y = AK^{\alpha}L^{1-\alpha}$, we obtain[30]

$$\frac{\sum_n Y_n - \sum_n Y_n^c}{\sum_n Y_n^c} = \sum_n \frac{Y_n^c}{\sum_i Y_i^c}\left(\frac{Y_n - Y_n^c}{Y_n^c}\right) = \sum_n \frac{Y_n^c}{\sum_i Y_i^c}\left[\left(\frac{L_n}{L_n^c}\right)^{1-\alpha} - 1\right], \tag{10}$$

where $Y_n^c$ is the counterfactual output level of city $n$ without agencies, $L_n^c$ is the counterfactual labor stock without agencies and $L_n$ is the labor stock with agencies. The above equation says that assuming fixed capital stock and productivity, aggregate output gains due to agencies can simply be calculated as a weighted average of the labor gains, weighted by the counterfactual output share for each city $\frac{Y_n^c}{\sum_i Y_i^c}$.

To measure the city output share, I calculate each city's ex ante share of national GDP using the Chinese City Statistical Yearbook. I take $\alpha = 0.3$ which follows from (i) the capital output ratio of manufacturing firms in the administrative tax survey dataset and (ii) running a regression of city level industrial output on industrial fixed assets and total labor input and obtaining the coefficient of 0.29 for the capital input[31]. To determine the low-skilled labor share $\beta$, I find that the payment for managers takes up around 25% of total labor payment among manufacturing firms in the tax survey dataset. Using this as a proxy for the share of low-skilled labor over total labor share $\beta/(1-\alpha)$, I

[30] It is easy to obtain similar result using $Y = AK^{\alpha}(L_U)^{\beta}(L_S)^{1-\alpha-\beta}$.

[31] Specifically, I obtain $\beta_1 = 0.29$ from the two-way fixed effects model $y_{nt} = \beta_1 K_{nt} + \beta_2 L_{nt} + \nu_n + \gamma_t + \varepsilon_{nt}$.

Table 8: Relative Output Gains in 2010

| | $Y = AK^{\alpha}L^{1-\alpha}$ | $Y = AK^{\alpha}(L_U)^{\beta}(L_S)^{1-\alpha-\beta}$ | | |
|---|---|---|---|---|
| | | $\beta = 0.2$ | $\beta = 0.4$ | $\beta = 0.6$ |
| Relative Output Gains | 3.15% | 2.30% | 2.79% | 3.32% |
| Agency Profit Share | 39.82% | | | |

**Notes:** This table reports the aggregate output gains *relative* to the no agency counterfactual, as well as the share of agency profits in aggregate output gains in 2010.

obtain a benchmark value of $\beta = 0.2$, but will also consider alternative levels of $\beta$ and explore how they affect our aggregate gains calculations.

Now the core task boils down to finding the counterfactual total labor supply in each city with and without agencies, i.e., $L_n$ and $L_n^c$. For this, we can first use the results in Section 3.2.1 to establish inter-regional labor reallocation with labor agencies, as well as the counterfactual labor flow without them. Given Table 3, it then follows that when setting $Connected_{mnt} = Connection_{mnt} = 0$, one can obtain the counterfactual labor flow from city $m$ to city $n$ without agencies at year $t$ simply as $\hat{Flow}_{mnt}^{counterfactual} = \hat{\lambda}_{mn} + \hat{\gamma}_t$, and the fitted value $\hat{Flow}_{mnt}$ by plugging in the agency network $\{Connection_{mn}\}_t$ for predicted labor flows with labor agencies.

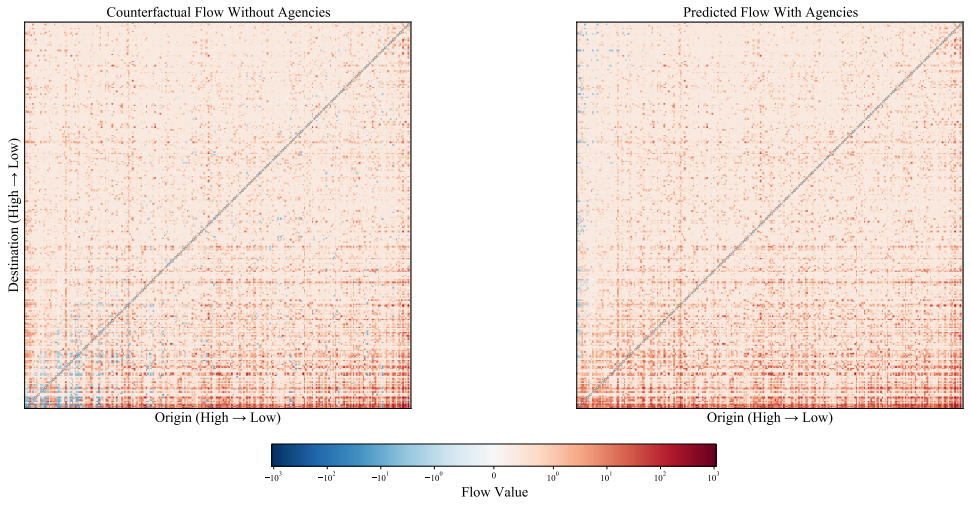
I contrast the total labor flow with agencies in 2010 with the counterfactual labor flow in the $N \times N$ city system (where $N$ is the total number of cities) in the top panel of Figure 8. The horizontal (vertical) axis shows origin (destination) cities ranked from the highest to the lowest in the ex ante average wage distribution. Hence the $i$-th column and $j$-th row pixel shows the total number of labor from from the $i$-th rank city to the $N + 1 - j$-th rank city in 2010 respectively with labor agencies (right) and in the counterfactual scenario (left). This also implies all pixels below the diagonal show labor flow from lower wage cities to higher wage cities and vice versa for the above diagonal pixels. The core message from this comparison is that labor agencies reinforce the underlying worker flows from low wage to high wage cities. First, one can see substantial flows of workers from cities across the ex ante wage distribution to the highest wage cities, and all the negative flows from low to high wage cities in the counterfactual scenarios get eliminated in the scenario with labor agencies. Second, there are negative labor flows in the top left corner of city system with agencies, which implies that with labor agencies, it is even less likely for workers to flow reversely, i.e., from ex ante high wage to low wage cities. Furthermore, I calculate the total headcount of worker flows across the entire city system attributable to agencies is around 27 million[32], which is roughly two third of the total population of agency workers (37 million) reported by ACFTU in 2011. This confirms that the magnitude of our back-of-the-envelope calculations is reasonably reliable.

The middle and bottom panels of Figure 8 show such fitted and counterfactual labor flows separately for college labor and below college labor. Similar patterns emerge: for both college and below college population, labor agencies reinforce the flows to the highest wage cities, and all the (counterfactual) negative flows from low to high wage cities get eliminated under labor agencies. Furthermore, compared to college labor, labor agencies lead to a stronger negative effects among below college labor for the reverse labor flow (i.e., from high to low wage cities). This is consistent with the predominant role of labor agencies in organizing low-skilled labor across regions.

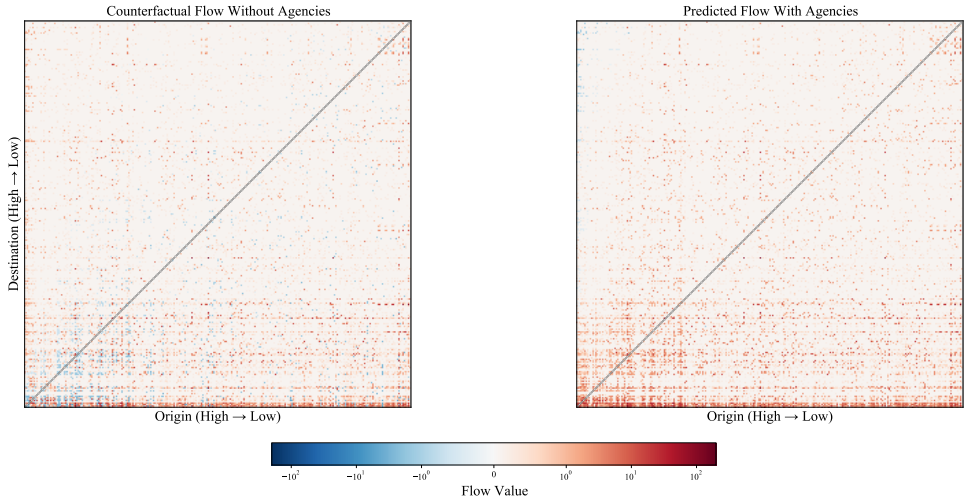
Using the above counterfactual labor flow results and ex ante distribution of city population, we can now calculate the aggregate output gains with labor agencies using equation (10). I report the

[32]Specifically, I sum up all predicted values of $\hat{Flow}_{mnt}$ and $\hat{Flow}_{mnt}^{counterfactual}$ over all city pairs in 2010, and the difference in these two sums is the total worker flow attributable to agencies.

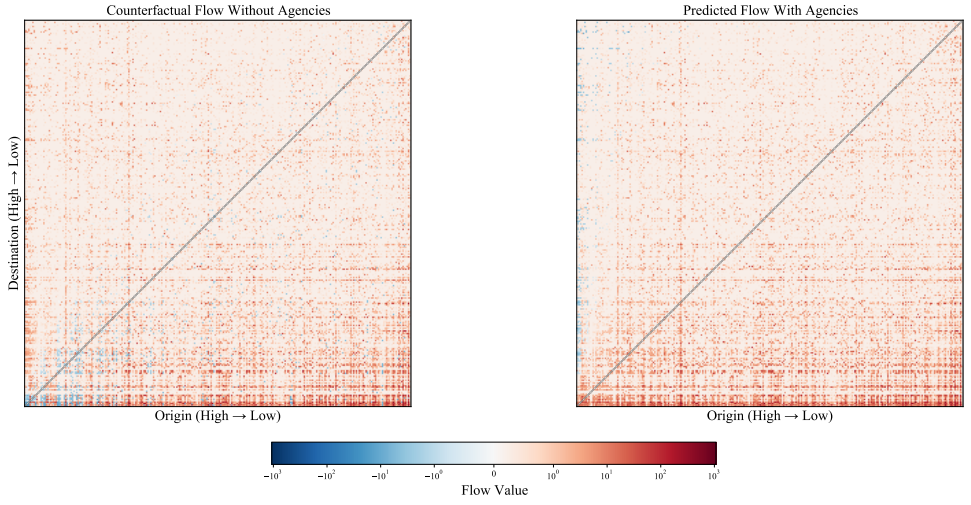
FIGURE 8: INTER-CITY LABOR REALLOCATION




(A) ALL




(B) COLLEGE




(C) BELOW COLLEGE

*Notes:* This figure shows the counterfactual labor flow (left), $\hat{Flow}_{mn}^{counterfactual}$, and the fitted value of labor flow with agencies (right), $\hat{Flow}_{mn}$ in the $N \times N$ city system in 2010, respectively for total labor flow (top panel), individuals with a college or above degree (middle panel) and those without a college degree (bottom panel). The horizontal (vertical) axis shows the origin (destination) city, ordered from left (bottom) to right (top) according to their ex ante average wage. Therefore the $i$-th column and $j$-th row pixel shows the total number of labor flow from the city with $i$-th rank in the ex ante average wage distribution to the city with $N+1-j$-th rank in the ex ante average wage distribution.

aggregate output gains by labor agencies *relative* to the no agency counterfactual in Table 8. Using total labor flows, I find that the existence of labor agencies across cities in 2010 can contribute to 3.15% aggregate output gains in China relative to the counterfactual scenario without agencies. I also find a 2.30-3.32% aggregate gains using the skill specific production function with different values of $\beta$, where higher values of $\beta$ yields higher estimates for the relative aggregate gains. This is intuitive because we previously found a stronger effect of labor agencies to move low-skilled workers across cities, and as low-skilled labor becomes more important in the production function (i.e., higher $\beta$), one can expect agencies to generate a larger aggregate impact.

These estimates should be interpreted with two caveats. First it only captures the short-run output gains from labor reallocation, holding city-level capital stocks fixed. In the model in Section 4, agency expansion also accelerates capital accumulation in destination cities, which implies larger medium to long run output gains once capital is allowed to adjust. Second, the numbers reported in Table 8 are output gains rather than surplus gains. They therefore do not subtract the costs of intermediation and relocating workers across regions. The model section incorporates this margin explicitly by allowing agencies to incur convex intermediation costs, providing a disciplined way to distinguish gross output gains from net surplus gains.

Moreover, we can have a back-of-the-envelope estimate of how much of such aggregate output gains are captured by labor agencies as profits. To calculate agency profits in reduced-form, I first use the subsample of agencies from the administrative tax survey dataset to run a regression of agency headquarter profits per worker on ex ante maximum wage difference between cities that each agency group later expanded into using a specification similar to that in Table 2[33]. I then obtain that one unit[34] increase in the ex ante wage gap translates into 1.65 unit increase in agency profits per worker, but one extra competing agency group between the corresponding city pair reduces profits by 0.58 unit. This corresponds to the fact that competition among agencies may reduce the profit margins that they can capture when contracting with firms and workers. Furthermore, one percent increase in the distance among branches with the largest wage gap is associated with 0.85 reduction in the agency profits, which is associated with the costs in reallocating workers across space.

We can then calculate the per worker profits for a given agency group operating between cities $m$ and $n$ as $1.65 \times \Delta Wage_{mn}^{pre} - 0.58 Connection_{mn} - 0.85 Distance_{mn}$. Multiplying this by the total number of additional labor flow due to agencies $\hat{Flow}_{mn} - \hat{Flow}_{mn}^{counterfactual}$, as well as the number of agency groups that connect the cities $Connection_{mn}$, I can obtain total profits of the agency sector when reallocating labor between cities $m$ and $n$. Aggregating such agency profits over the entire $N \times N$ city system, I find that total agency profits takes up around 40% of aggregate output gains[35], leaving the remaining 60% of them to firms and workers. The fact that around 40% of the output gains from labor reallocation are captured by the intermediaries is however consistent with the previous empirical facts that there might be limited competition among agencies (i.e., many of the city pairs are connected via only one agency group, and the number of competing agencies is associated with

[33] Specifically, I run

$$Profits_g = \beta_1 \left|\max \Delta Wage_g^{pre}\right| + \beta_2 Number\ of\ Agency\ Groups_g + \beta_3 Distance_g + \epsilon_g,$$

where $|\max \Delta Wage^{pre}|$ is the maximum ex ante (i.e., 2007) wage difference among cities that the agency group later expanded into during 2008-2018. Number of Agency Groups is the total number of agency groups that had established branches in the same cities that exhibit such maximum wage gap. $Distance_g$ is the geographical distance between the cities that exhibit such maximum wage gap

[34] The units for both profits and city wage gap are 1k RMB.

[35] We take the total GDP in 2005 from the Chinese City Statistical Yearbook as the no agency output, hence the aggregate output gain is simply 3.15% of such total GDP.

lower profits).

I can also run counterfactual analysis for aggregate output gains by incorporating the expansion of labor agencies after 2010. I keep the no agency counterfactuals the same as in 2010, i.e., $\hat{Flow}_{mnt}^{counterfactual} = \hat{\lambda}_{mn} + \hat{\gamma}_t$, but I calculate the predicted values of flows with agencies by taking the future expansion of agency network into account. Therefore the core question we ask in this exercise is how much additional aggregate output gains can be delivered in 2010, if one is to have more agencies across cities as in 2011 - 2018. I report these relative output gains calculations in Table A.13. Labor agencies continue to expand to more cities after 2010, and as one would expect, there will be larger gains in aggregate output, which ranges from around 8% to 20% given the agencies network respectively in 2011 and 2018. With the skill specific aggregate production function, I find similar levels of output gains: for example, using the benchmark value of $\beta = 0.2$, I find that the agencies network as in 2018 can deliver 28% aggregate output gains, and these aggregate gains are all increasing as agencies continue to expand over time.

The fact that labor market intermediaries can appropriate 40% of the aggregate gains as profits speaks to the importance of agencies competition in determining the functional distribution of surplus gains among economic groups. This also implies that one need to account for the endogenous agency dynamics into the process of reducing misallocations, and, importantly, how it interacts with the dynamic path of regional convergence. A series of interesting questions emerge: for example, how will agencies affect the dynamics of labor transition and capital accumulation in the long run, how much can agencies benefit along the transition path, and how persistent the aggregate gains can be even after agencies exit. I turn to these issues now.

# 4 A Neoclassical Growth Model with Labor Agencies

We now develop a neoclassical growth model with spatial mobility frictions and endogenous entry and exit of labor market intermediaries. The empirical analysis and the back-of-the-envelope calculations in the previous section quantify the short-run effects of agency networks on labor reallocation and aggregate output, taking regional capital stocks as fixed. This section instead studies the longer-run implications of agency-mediated labor reallocation when regional capital stocks adjust over time.

The model serves two purposes. First, it studies how temporary agency activity shapes the transition from less productive to more productive regions. The model features a dynamic amplification effect: agencies do not change the efficient steady state, as they operate only when regional wage disparities create arbitrage opportunities and exit once these opportunities disappear. Yet by reallocating labor along the transition path, they can amplify aggregate gains through regional capital accumulation. Second, the model examines what determines the distribution of surplus gains along the transition path of regional convergence. Motivated by the empirical finding that agencies capture a substantial share of the gains from intermediation as profits, the model asks whether this surplus division is merely distributional, or whether it also has efficiency consequences because it reflects underlying agency market power and competition.

## 4.1 The Model

### 4.1.1 The Basic Set Up

We consider two regions $r \in \{R, P\}$ with productivity parameters $A_R > A_P$, and there are two types of economic agents: workers and entrepreneurs. Furthermore, we assume two skill levels for workers $s \in \{U, S\}$ where $U$ stands for low-skilled and $S$ is skilled. The total stock of type $s$ labor in the

economy is $\overline{L}_s$. Furthermore, we assume firms/entrepreneurs in each region owns a fixed amount of land, $\overline{D}_r$, and there is no competitive land (rental) market. This is to ensure an interior solution for the steady state with constant capital stock in both regions. Each region has a production technology given by

$$y_r = A_r \left(K_r\right)^{\alpha} \left(L_{Ur}^{\rho} + L_{Sr}^{\rho}\right)^{\frac{\beta}{\rho}} \overline{D}_r^{1-\alpha-\beta} \tag{11}$$

where $K_r$ is physical capital, and $L_{sr}$ is labor input of skill type $s$ in region $r$. The elasticity of substitution between the two skill types is $\frac{1}{1-\rho}$. Let us assume that physical capital is immobile, as we interpret physical capital as buildings and structures as in e.g., Kleinman et al. (2023). This implies that entrepreneurs rely on their own capital stock for production. We also assume that firms can only hire labor locally, so there will be region specific wage for each skill level, $w_r^s$. Therefore for each period given their capital stock, entrepreneurs will choose labor inputs to maximize profits. Firm profit therefore becomes

$$\pi_r \left(K_r\right) = \max_{L_{Ur}, L_{Sr}} A_r \left(K_r\right)^{\alpha} \left(L_{Ur}^{\rho} + L_{Sr}^{\rho}\right)^{\frac{\beta}{\rho}} \overline{D}_r^{1-\alpha-\beta} - w_r^U L_{Ur} - w_r^S L_{Sr}. \tag{12}$$

**Capital Accumulation** The inter-temporal consumption-investment decision of entrepreneurs can be formulated as follows: taking initial capital stock $K_{r0}$ as given, entrepreneurs in region $r$ choose the consumption path $\{c_{rt}\}_{t\geq 0}$ to maximize lifetime utility

$$\sum_{t=0}^{\infty} \gamma^t \frac{c_{rt}^{1-\theta} - 1}{1-\theta} \tag{13}$$

subject to the budget constraint:

$$K_{rt+1} = \pi_r \left(K_{rt}\right) + (1-\delta) K_{rt} - c_{rt} \tag{14}$$

where $\delta \in (0,1)$ is capital depreciation rate and $\gamma \in (0,1)$ is the discount factor. The Euler equation is

$$\begin{aligned} \left(\frac{c_{r,t+1}}{c_{rt}}\right)^{\theta} &= \gamma \left[\pi_r'(K_{rt+1}) + 1 - \delta\right] \\ &= \gamma \left[\alpha A_r K_{rt+1}^{\alpha-1} \left((L_{Urt+1}^*)^{\rho} + (L_{Srt+1}^*)^{\rho}\right)^{\frac{\beta}{\rho}} \overline{D}_r^{1-\alpha-\beta} + 1 - \delta\right] \end{aligned}$$

where the second equality follows from envelope theorem, and $L_{srt}^*$ denotes the optimal labor input determined in (12).

**Imperfect Labor Mobility** For simplicity we assume workers cannot save; therefore for each period they will migrate to whichever region that gives them higher wages and immediately consume their earnings, subject to the spatial mobility barriers implicitly modeled below. We follow Gennaioli et al. (2014) for a tractable way to introduce mobility barriers into the neoclassical growth model[36]: if there is infinite mobility cost, then each region's labor stock at period $t$, $L_{srt}$, simply equals to the level at the previous period, i.e., $L_{srt-1}$, i.e., workers just stay at where they were in the previous period. On the other hand, if there is perfect mobility, workers will flow freely so that the skill specific wage rate

[36] In contrast with the current models on rural-urban migration which often feature seasonality and taste shocks as in e.g., Lagakos et al. (2023), this paper focuses on labor mobility across cities/regions, and with this simple way of introducing mobility barriers, I concentrate on how agency activities interact with the transition dynamics of regional convergence.

at period $t$ will equalize across regions. If this is the case, the supply of type $s$ workers at period $t$ will simply be proportional to regional capital stock and productivity parameters, i.e.,

$$\frac{\hat{L}_{URt}}{\hat{L}_{UPt}} = \frac{\hat{L}_{SRt}}{\hat{L}_{SPt}} = \left( \frac{A_R K_{Rt}^{\alpha} \overline{D}_R^{1-\alpha-\beta}}{A_P K_{Pt}^{\alpha} \overline{D}_P^{1-\alpha-\beta}} \right)^{\frac{1}{1-\beta}}. \tag{15}$$

Therefore the (hypothetical) allocation of type $s$ labor for region $r$ becomes

$$\hat{L}_{srt} = \frac{\left( A_r K_{rt}^{\alpha} \overline{D}_r^{1-\alpha-\beta} \right)^{\frac{1}{1-\beta}}}{\sum_q \left( A_q K_{q,t}^{\alpha} \overline{D}_q^{1-\alpha-\beta} \right)^{\frac{1}{1-\beta}}} \overline{L}_s \equiv Q_{rt} \overline{L}_s. \tag{16}$$

Then for intermediate level of mobility barriers, we assume that the total stock of type $s$ worker in region $r$ evolves according to the following "sluggish" updating rule:

$$L_{srt} = \nu_{st} \left( L_{srt-1} \right)^{\tau_s} \left( \hat{L}_{srt} \right)^{1-\tau_s}, \tag{17}$$

where the normalization factor is $\nu_{st} \equiv \frac{\overline{L}_s}{\sum_q (L_{sqt-1})^{\tau_s} \left( \hat{L}_{sqt} \right)^{1-\tau_s}}$, which makes sure that $\sum_q L_{sqt} = \overline{L}_s$ for all $t$. In this model, the role of spatial mobility barriers is to delay the adjustment to the efficient allocation of labor. Furthermore, we allow for skill specific degree of mobility barriers $\tau_s$ across the two regions, which captures the fact that better educated workers may have more information and/or connections with job opportunities in rich regions.

Before we introduce labor agencies, we can characterize the steady state of the economy and summarize the results here:

**Proposition 1** *The steady state of the economy is unique and efficient such that despite the mobility barriers, the long-run capital and labor allocations are proportional to regional productivity parameters:*

$$\frac{K_R^{**}}{K_P^{**}} = \frac{L_{UR}^{**}}{L_{UP}^{**}} = \frac{L_{SR}^{**}}{L_{SP}^{**}} = \left( \frac{A_R}{A_P} \right)^{\frac{1}{1-\alpha-\beta}} \frac{\overline{D}_R}{\overline{D}_P}. \tag{18}$$

This result parallels the previous arguments in the literature such that market frictions may not be handicap to long-run outcomes, as agents can gradually accumulate physical or human capital overtime to overcome these constraints and achieve efficient outcomes (see, for example, Banerjee and Moll (2010), Ghatak (2015)). This allows us to concentrate only on the role of market intermediation in speeding up the convergence process (Barro et al. (1995)). We now model how labor recruitment agencies can impact the convergence process.

### 4.1.2 Labor Agencies

We assume that agencies can *temporarily* move *low-skilled* workers across regions [37]: suppose at period $t$, labor agencies decide to hire $L_{at}$ of type $U$ workers in region $P$ at wage $w_{at}$ (which will be determined below), and then move them to region $R$ so that they can temporarily join the workforce at region $R$[38]. Then at the end of period $t$, with probability $q$, these workers can successfully settle

[37] Since our focus here is to look at how agencies speed up regional convergence in the long run, we abstract away from equilibrium unemployment.

[38] By "temporary," I mean agency-mediated labor reallocation that need not involve permanent migration or a change in registered residence. Direct evidence on the duration of individual agency assignments and their subsequent migration

down in region $R$ and become a permanent worker there[39]; with probability $1-q$, agency workers will return back to region $P$. $q$ may capture the *Hukou* quota for migrant workers, or more generally the administrative or cultural barriers for agency workers from formally integrating into the new region even though they have temporarily worked there.

**Timing** The timing within each period is summarized as follows:

1. At the beginning of period $t$, the permanent stocks of skilled and unskilled labor in each region, $\{L_{srt}\}_{s,r}$, are determined by the sluggish mobility rules. For unskilled workers, this includes the fraction $qL_{a,t-1}$ of agency workers from period $t-1$ who successfully settled in region $R$.

2. Given the permanent labor stocks $\{L_{srt}\}_{s,r}$ and capital stocks $\{K_{rt}\}_r$, agencies decide whether to operate. If they operate, agencies bargain with firms in region $R$ over the number of agency workers $L_{at}$ and the price $p_{at}$ paid by firms.

3. Agencies recruit $L_{at}$ low-skilled workers from region $P$ at wage $w_{at}$ and temporarily place them in region $R$ for production during period $t$. These workers enter production in region $R$ as agency labor but are not part of the permanent labor stock of region $R$ during period $t$.

4. At the end of period $t$, a fraction $q$ of agency workers successfully settles in region $R$ and becomes part of its permanent low-skilled labor force in period $t+1$. The remaining fraction $1-q$ returns to region $P$.

5. Entrepreneurs choose consumption and investment, which determines the next-period capital stocks $\{K_{r,t+1}\}_r$.

**Production with Agency Labor** Following Felix and Wong (2024), we assume that local unskilled labor and agency workers are imperfect substitutes. Hence the regional production function with agency workers is

$$y_r^a = A_r \left(K_r\right)^{\alpha} \left(H_{Ur}^{\rho} + L_{Sr}^{\rho}\right)^{\frac{\beta}{\rho}} \overline{D}_r^{1-\alpha-\beta}, \tag{19}$$

where the total effective unskilled labor is

$$H_{Ur} = \left[L_{Ur}^{\eta} + \zeta L_a^{\eta}\right]^{\frac{1}{\eta}}, \tag{20}$$

where $\zeta \leq 1$. Therefore entrepreneurs in region $R$'s static profit maximization problem becomes the following: taking the number of low-skilled workers $L_a$ introduced by agencies and the price $p_a$ as given, they choose $\{L_{UR}, L_{SR}\}$ to maximize per-period profit:

$$\tilde{\pi}_R(K_R, p_a, L_a) = \max_{L_{UR}, L_{SR}} A_R \left(K_R\right)^{\alpha} \left(H_{UR}^{\rho} + L_{SR}^{\rho}\right)^{\frac{\beta}{\rho}} \overline{D}_R^{1-\alpha-\beta} - w_R^U L_{UR} - w_R^S L_{SR} - p_a L_a. \tag{21}$$

where $p_a$ is the price that firms in region $R$ pay to the agencies for each unskilled worker they obtain. We denote the output level evaluated at the optimal choice of $\{L_{UR}, L_{SR}\}$ in (21) as $\tilde{y}_R(L_a)$.

patterns is limited, since labor surveys following agency workers across cities over long horizons are rare. The empirical evidence earlier in this paper is nevertheless consistent with this interpretation: agency connections generate sizable cross-city labor flows but much smaller changes in cross-city *Hukou* population gaps, suggesting that agencies mainly facilitate labor mobility without inducing comparable changes in permanent registered residence.

[39] They will be subject to the same mobility barriers described above should they wish to move back to region $P$.

**Agencies and Mobility** In Appendix B.4, I derive the regional allocation of workers under the (hypothetical) frictionless mobility case, i.e., $\hat{L}_{sr}$, taking $L_{at}$ as given. Furthermore, at the end of each period, $qL_{at}$ of unskilled worker (originally from $P$) will become a permanent resident in the agency destination region $R$. The dynamics of low-skilled labor mobility in region $R$ becomes

$$L_{URt} = \nu_{Ut} \left(L_{URt-1} + qL_{at-1}\right)^{\tau_U} \left(\hat{L}_{URt}\right)^{1-\tau_U}, \tag{22}$$

and the dynamics of low-skilled labor in region $P$ is[40]

$$L_{UPt} = \nu_{Ut} \left(L_{UPt-1} - qL_{at-1}\right)^{\tau_U} \left(\hat{L}_{UPt}\right)^{1-\tau_U}. \tag{23}$$

Here $\nu_{Ut}$ is defined analogously to what we have before, adjusted for the permanent moves $qL_{at-1}$. The dynamics of skilled labor mobility is the same as (17). The temporariness of such labor allocation via agencies also implies that when the agency ceases to be active, the $(1-q)L_a$ amount of agency workers will have to immediately return to their origin region and are again subject to the same mobility dynamics specified before.

We now describe how the number of agency workers $L_{at}$, their wage $w_{at}$, and $p_{at}$ are determined. For now let us temporarily drop the $t$ subscript to model the static optimization of firms and agencies.

**Selection into Agencies and Labor Market Power** Unskilled workers $i$ in region $P$ will join agencies if and only if[41]

$$w_a \upsilon_i \geq w_P^U, \tag{24}$$

where $\upsilon_i$ is an idiosyncratic taste shock which captures the perceived utility of traveling to region $R$, agency job amenities, etc. The cumulative density function of $\upsilon_i$ is $G(\upsilon_i)$. Hence the supply of agency workers is

$$L_a = \left[1 - G\left(\frac{w_P^U}{w_a}\right)\right] L_{UP}. \tag{25}$$

We can invert this function and obtain $w_a$ as a function of $L_a$, i.e., $w_a(L_a)$.

We assume that agencies have labor market power as in Amodio et al. (2025) and Kline (2025) such that the profit of the agencies is[42]

$$W(p_a, L_a) = \left(p_a - w_a(L_a)\right) L_a - \phi(L_a) - F \tag{26}$$

where $\phi(L_a)$ is a convex cost function of training and transporting $L_a$ low-skilled workers from region $P$ to region $R$.[43] This captures the capacity constraints of agencies, and/or the diminishing returns

[40] We assume that $q$ is low such that $L_{UPt} > qL_{at}$ for all $t$.

[41] The timing is as follows: at the beginning of each period, workers decide to permanently move across regions according to (22), (23), and (17), taking $L_{at}$ as given. For unskilled workers who decide to stay at region $P$, they will then have an option of joining an agency.

[42] As mentioned in Section 2.1, a well-functioning system of labor unions is largely absent in the empirical context of this paper; therefore the "right-to-manage" bargaining models between workers/labor unions and agencies (Harris (1982), Holmes and Snider (2011), Felix and Wong (2024)) seem less plausible. We also abstract away from information asymmetry problems in agency activities. See a recent paper by Fernando and Singh (2025) which demonstrates how the government's rating program can increase intermediaries' investment in screening efforts and job placement quality for international migrants.

[43] We are agnostic about the specific forms of search frictions, and we assume that the agencies technology has some comparative advantages in temporarily reallocating workers across regions subject to the convex cost function. See, for example, Heise and Porzio (2023) which embed a search framework into a model with spatial frictions, and Bilal and Lhuillier (2025) which model how agencies mitigate search frictions. More broadly, the recent review by Gong et al. (2025) discusses the literature on market intermediaries and search equilibrium.

in recruitment/placement efforts as well as maintaining good reputation among firms and workers. $F$ is the fixed cost in agency operations.

**Bargaining between Firms and Agencies** Entrepreneurs in region $R$ and agencies then bargain over a contract which specifies $\{p_a, L_a\}$[44]:

$$\max_{p_a, L_a} W(p_a, L_a)^{\chi} \left[\tilde{\pi}_R\left(K, p_a, L_a\right) - \pi_R\left(K\right)\right]^{1-\chi}, \tag{27}$$

where $\chi$ represents the bargaining power of agencies relative to entrepreneurs, and $\pi_R(K)$ is the entrepreneurial profits without agencies as defined in equation (12). We set the outside option of agencies to be zero. One can also find similar bargaining structure between firms and agencies in recent models of labor outsourcing (for example, Felix and Wong (2024)). We can then solve the optimal $\{p_a^*, L_a^*\}$ and summarize the results as follows:

**Proposition 2** *The optimal number of agency workers is characterized by*

$$\tilde{y}_R'(L_a) = w_a(L_a)\left[1 + \frac{w_a'(L_a)L_a}{w_a(L_a)}\right] + \phi'(L_a^*). \tag{28}$$

*The price that firm pay to agency $p_a^*$ is characterized by*

$$(1-\chi)W(p_a^*, L_a^*) = \chi\left[\tilde{\pi}_R\left(K, p_a^*, L_a^*\right) - \pi_R\left(K\right)\right]. \tag{29}$$

Equation (28) says that the number of agency workers should be chosen in a way to balance between the marginal increase in region $R$'s output, on the one hand, and the marginal cost in incentivizing workers to join the agency by providing labor wage as well as the marginal cost in organizing agency workers, on the other. The labor market power of the agencies is captured by the elasticity term $\frac{w_a'(L_a)L_a}{w_a(L_a)}$, which gives rise to wage markdown in the classic monopsony framework. This also implies that $L_a$ is chosen in a way such that agency workers wage $w_a$ is lower than the marginal gains in production efficiency $\frac{\partial \tilde{\pi}_R(K, p_a, L_a)}{\partial L_a} - \phi'(L_a^*)$.

Equation (28) also suggests that the bargaining power of agencies $\chi$ doesn't affect the intensive margins of labor reallocation $L_a$. However it does affect how the price of agency workers that firms pay, and in turn the distribution of surplus among agencies and firms. Equation (29) implies that the profits that agencies obtain will take up $\chi$ share of the total (static contract level) surplus generated from labor market intermediation: $\tilde{\pi}_R\left(K, p_a^*, L_a^*\right) - \pi_R\left(K\right) + W(p_a^*, L_a^*)$.

Agencies decide on whether to operate or not at the beginning of each period, i.e., they will operate if and only if $W(p_a^*, L_a^*) \geq 0$.

**Capital Accumulation** For the inter-temporal consumption-investment decision, the budget constraint of the entrepreneurs will be the same as equation (14) except that we replace $\pi_r(K)$ by $\tilde{\pi}_R(K, p_a, L_a^*)$ for region $R$. Note also that $\tilde{\pi}_R = \pi_R$ when labor agencies are not active; for region $P$, $\tilde{\pi}_P = \pi_P$. The Euler equation is again

$$\left(\frac{c_{rt+1}}{c_{rt}}\right)^{\theta} = \gamma\left[\tilde{\pi}_r'(K_{rt+1}) + 1 - \delta\right]$$

[44] We assume that bargaining takes place before the fixed cost is incurred, hence we avoid the hold up problem in contract negotiations.

**Market Clearing Conditions** Regional wages clear the local labor markets. It is useful to note that the supply of unskilled labor in region $P$ is $G\left(\frac{w_{UPt}}{w_{at}}\right) L_{UPt}$. First order conditions from (12) gives the optimal labor demand in $P$. Together with (25) and (28), these equations jointly determine $w_{at}$, $L_{at}$, $w_{Pt}^U$, and $w_{Pt}^S$. For region $R$, the first order conditions from (21) and the local supply of workers $L_{sRt}$ determine $w_{Rt}^s$.

Moreover, we can intuitively see how intermediation affects the speed of convergence: they bring labor across regions to arbitrage the wage disparities more quickly than workers' mobility by themselves; they also bring efficiency gains to entrepreneurs and indirectly speed up the capital accumulation process by raising marginal products of capital (since capital and labor are complements). Importantly, this implies that even if agencies exit because of limited arbitrage opportunities, we still expect the aggregate output gains along the transition path to be persistent because the rich region now has a higher stock of capital compared to the no agencies counterfactual. Furthermore, when agencies exit because of limited arbitrage opportunities, agency workers temporarily working in region $R$ will have to go back to region $P$ with probability $1-q$, leading to a sharp drop in low-skilled labor available to region $R$, which will then slowly recover subject to mobility barriers.

Also we can show that the economy will converge to the same efficient steady state as characterized in Proposition 1. The core logic is that since marginal product of capital and labor input will be eventually equalized across regions, there is no room for agencies to arbitrage so in steady state we must have $W < 0$. I provide the proof in Appendix Section B.6. This suggests that labor agencies will only affect the transition dynamics to the steady state, and we will now turn to this with the quantitative results. I will first describe how parameter values are calibrated to empirical moments.

## 4.2 Calibration

For the factor share parameters in the production function, we take capital share $\alpha$ to be 0.3 to match the capital output ratio of manufacturing firms in the administrative tax survey dataset (as in Section 3.3). Consistent with Hsieh and Klenow (2009), we take the labor share $\beta$ to be 0.5 in China. I follow Fang and Herrendorf (2021) to calibrate $\rho = 0.3$ such that the elasticity of substitution between skilled and unskilled labor is around 1.42. For the elasticity of substitution between regular and agency workers, Micco and Munoz (2024) obtain a model-based estimates of 4.28, which translates into $\eta = 0.8$. I calibrate $\zeta = 0.7$ to match the earnings ratio of below college agency workers over directly hired workers in the CLDS dataset.

For the regional productivity parameters $A_R$ and $A_P$, I use the Chinese City Statistical Yearbook dataset where I run city level industrial output on industrial fixed assets and total labor input, and I collect the city fixed effects to measure city level industrial productivity. Then I calculate average productivity respectively for cities in the top 10% of the 2005 average wage distribution, and those in the bottom 90%. These correspond to region $R$ and region $P$ in the model. I normalize $A_R = 1$, and obtain $A_P = 0.9$.

Then using the 2005 population census data, I find that among respondents aged between 25-55, the population with at least a college degree is around 10% of population without a college degree. I therefore take the total stock of labor $\overline{L}_U = 1$, and $\overline{L}_S = 0.1$. Furthermore, I find that the 2005 top 10% wage cities take up around 15% of all below college respondents, and 45% of all respondents with a college degree. This implies $L_{UP0} = 0.85$, $L_{UR0} = 0.15$, $L_{SP0} = 0.055$, $L_{SR0} = 0.045$. For initial capital stock, I find from the Chinese City Statistical Yearbook data that the average industrial fixed asset among cities with bottom 90% average wage in 2005 is around 20% of that among cities with top 10% average wage. I therefore take $K_{R0} = 1$ and $K_{P0} = 0.2$. I will simulate the model over time

periods $t \in \{0, 1, ..., 150\}$, where I assume that the economy will converge to the steady state at the 150th period.

We assume the following simple quadratic form for the cost function of agencies:

$$\phi(L) = \frac{1}{2}\phi L^2. \tag{30}$$

To allow for sufficient labor reallocation via agencies in the simulation, I assume $\phi = 0.1$ and $F = 0.1$. As discussed in Kline (2025), I assume a power function CDF for $G(\upsilon)$:

$$G(\upsilon) = \upsilon^{\psi}, \tag{31}$$

and consider a benchmark value $\psi = 4$. This functional form implies that the monopsony wedge in equation 28 is

$$\frac{w_a'(L_a)L_a}{w_a(L_a)} = \frac{L_a}{\psi(L_{UP} - L_a)}, \tag{32}$$

and therefore lower levels of $\psi$ implies stronger wage markdown. In the quantitative analysis, I will vary $\psi \in \{1, 2, 4\}$ to capture different degrees of labor market power of agencies in the labor origin cities.

We assume capital depreciation rate $\delta = 5\%$, take $\theta = 2$ to capture a reasonable level of risk aversion in preferences, and take $\gamma = 0.98$ for the patience parameter. For the mobility barriers $\tau_s$, notice first that Gennaioli et al. (2014) consider the values between 0.87 and 0.99 implied by cross-country regressions of growth convergence. For the results below, we assume the level of mobility barriers such that $\tau_U = 0.95$ and $\tau_S = 0.85$ to guarantee that it takes times for regions to converge to the steady state. This also implies that skilled workers face lower mobility barriers than the low-skilled workers, as they tend to have better information or they are subject to relaxed *hukou* barriers of movement to richer cities. We take $q = 0.06$ as a benchmark value to match the fact that the probability of successful applications among (floating) migrants for obtaining *hukou* in Beijing is around 5.6% in 2019. Note that this might represent the hardest case for migrants to obtain *hukou* in Chinese cities. In the policy counterfactual exercises, we will also consider a higher level of $q \in \{0.06, 0.09, 0.12\}$ to explore how this affects the transition dynamics.

Lastly, in the benchmark result below, we consider the bargaining power parameter $\chi = 0.5$, which says that agency will capture 50% of the total surplus from intermediation. But we will later consider alternative cases of $\chi \in \{0.1, 0.3, 0.5\}$ to capture different degrees of competition among agencies where the quantitative results may vary.

### 4.3 Benchmark Results

Focusing on bargaining power $\chi = 0.5$, Figure 9 presents the transition dynamics of factor prices, capital accumulation, as well as aggregate surplus to the steady state.

We can see how unskilled wages evolves along the transition path in the panel (a) of Figure 9, which highlights the role of labor agencies by making a contrast with a counterfactual dynamic path where no agencies are allowed to operate. The dashed curves are such counterfactual path for the two regions: there is large wage disparity in the initial periods, but regional wages gradually converge to each other as labor slowly moves across regions subject to the mobility barriers. When we allow labor agencies to operate (which I show using the solid curves), they drastically bring down low-skilled wages in region $R$ by introducing $L_a$ agency workers, and they also raises region $P$'s unskilled wages. This is consistent with the empirical results that the expansion of agency network is associated with

reduction in low-skilled wage disparities among pairs of cities. The low-skilled wage gap between the two regions continues to decrease while agencies are active and workers move across regions subject to mobility barriers. However, when agencies shut down at around the 40th period because of limited arbitrage opportunities from regional wage gap, $1-q$ of agency workers who are brought to region $R$ temporarily by agencies will now have to go back to region $P$, since the temporary labor reallocation is no perfect substitute for costly migration. Such temporariness of labor allocation via agencies leads to a sharp spike (drop) in region $R$ ($P$)'s low-skilled wage, which will then shrink over time as we are back to the case where workers cannot rely on agencies to transport them across regions.

I contrast the dynamic paths of capital accumulation and skill premium in panels (b) and (e) of Figure 9 with those in the counterfactual scenario when agencies are never allowed to operate. Region $R$ is absorbing low-skilled workers from region $P$ while accumulating capital, which is the core process of development as in Lewis (1954). The top right panel suggests that compared to the counterfactual scenario, the additional agency workers $L_a$ available to region $R$ raises the marginal returns of capital and therefore entrepreneurs' incentives to accumulate capital, consistent with our qualitative analysis before. This speaks to the wider development strategies that stress on keeping labor wages in the advanced sector low to speed up capital accumulation in the early stages of economic transition (see, for example, a recent argument by Itskhoki and Moll (2019)). On the other hand, there are reduced incentives for saving among entrepreneurs in region $P$ as they lose $L_a$ of low-skilled workers. In the bottom left panel, I find that labor agencies reduce skill premium (i.e., $w_r^S/w_r^U$) in region $P$ and raises the skill premium in region $R$. This directly follows from the fact that reallocating low-skilled labor from $P$ to $R$ reduces the low-skilled wage in $R$ but raises the low-skilled wage in $P$.

Panel (f) shows how the number of agency workers $L_a$, agency worker wage $w_a$, as well as the price that firms pay to agencies $p_a$ evolves along the transition path. Agencies are active for the first 38 periods and shut down thereafter. They moves around 30% of the low-skilled workers from region $P$ to region $R$ in the initial periods, but then $L_a$ gradually declines as wage gap shrinks and workers migrate. Bargaining power $\chi = 0.5$ allows the agency to charge $p_a$ around $1-2$, whereas the agency pays $0.5-1.5$ to workers from region $P$. Notably, both $w_a$ and $p_a$ increase over time since the local unskilled wage in region $P$ goes up and it becomes more and more costly for agencies to hire workers.

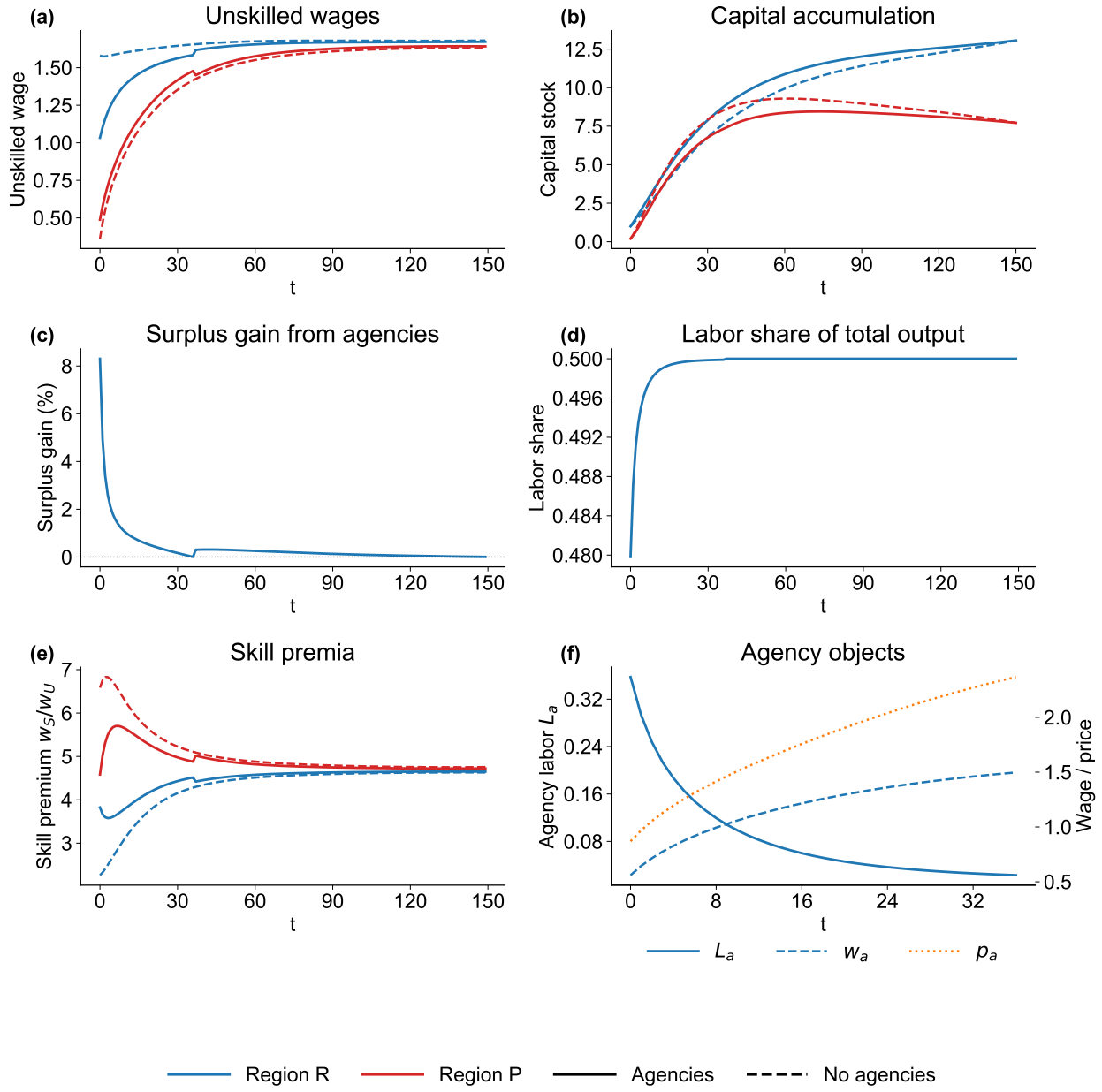
The middle panels (c) and (d) show the aggregate and distributional impact of labor reallocation via agencies. We can define aggregate surplus of the economy as

$$\sum_r y_r - \phi\left(L_a\right) - F, \tag{33}$$

i.e., the aggregate output minus the cost of moving agency workers across regions. I plot the dynamics of aggregate surplus gains in panel (c) of Figure 9. Agencies promote surplus gains (compared to the no agency counterfactual) in the initial periods, and the largest gain (8%) can be found in the first period when agencies reallocate the highest amount of workers across regions because of the largest arbitrage opportunities. However, such surplus gains diminish over time with reduced wage disparities and workers gradually move across regions. Importantly, the aggregate surplus gains are persistent even after agencies exit at the 38th period: because of temporary labor reallocations by agencies, region $R$ gets on faster track of capital accumulation and permanent labor transition, which makes the path of aggregate surplus always higher than that without agencies. In panel (d) of Figure 9, I plot the share of total labor income in aggregate output[45], which increases over time until it reaches

[45] This is defined as $(\sum_r \sum_s w_r^s L_{sr} + w_a L_a)/\sum_r y_r$.

FIGURE 9: BENCHMARK SCENARIO




*Notes:* The figure shows the transition dynamics of factor prices, capital accumulation, as well as aggregate surplus with $\chi = 0.5$. In the top left panel (low-skilled wages), the top right panel (capital allocations and skill premium), and the bottom left panel (skill premium $w_r^S/w_r^U$), the solid curves are for the benchmark scenario where $\chi = 0.5$, and the dashed curves are the counterfactual scenarios where no agencies are allowed to operate. The blue curves are for region $R$ and red curves are for region $P$. In the middle left panel, aggregate surplus with agencies is calculated as $\sum_r y_r - \frac{1}{2}\phi L_a^2 - F$, and I then report $(Surplus - Surplus^{no\ LA})/Surplus^{no\ LA}$. In the middle right panel, I show the share of total labor income in aggregate output, defined as $(\sum_r \sum_s w_r^s L_{sr} + w_a L_a)/\sum_r y_r$. In the bottom right panel, I show how the amount of agency workers $L_a$ (blue solid curve), agency worker wages $w_a$ (blue dashed curve), as well as the price that firms pay to agencies $p_a$ (red dashed curve) evolves over time (while agencies are active).

0.5 when agencies exit.

Taken together, these results demonstrate that initial misallocations can have dynamic implications: the arbitrage opportunities from large wage disparities due to labor mobility barriers incentivize agencies to bring workers from low wage to high wage regions and earn profits. This immediately brings down regional wage gap and generate large surplus gains in the initial periods. Because of the impact of temporary labor reallocations on capital accumulation and permanent labor transition, such surplus gains are also likely to be persistent even after agencies exit. We will now discuss how different scenarios of the agency sector affect the dynamic path of regional convergence and the distribution of surplus gains.

## 4.4 Policy Counterfactuals

We explore how different levels of bargaining power of agencies ($\chi$), degrees of labor market power of agencies ($\psi$), as well as the probability of transition of agency workers to permanent local worker ($q$) affect the gains from labor reallocation via agencies. We highlight the different channels through which the three parameters shape the aggregate gains and distributional implications. The empirical background for the policy counterfactuals in this section is the amendment to Labor Contract Law in 2013 that increases the entry barriers of labor dispatch agencies in China, which includes increased minimum capital requirements, additional certificates agencies need to obtain from local government, etc. The Law amendment leads to drop in agencies entries (as well as exit of some of the agencies with smaller scale), leading to potentially reduced competition among agencies. This implies higher bargaining power $\chi$ and higher labor market power (lower $\psi$). Furthermore, the discussion on $q$ will shed light on how policies such as lowering *hukou* barriers or facilitating agency workers' transition to permanent positions interacts with agency activities along the convergence path.

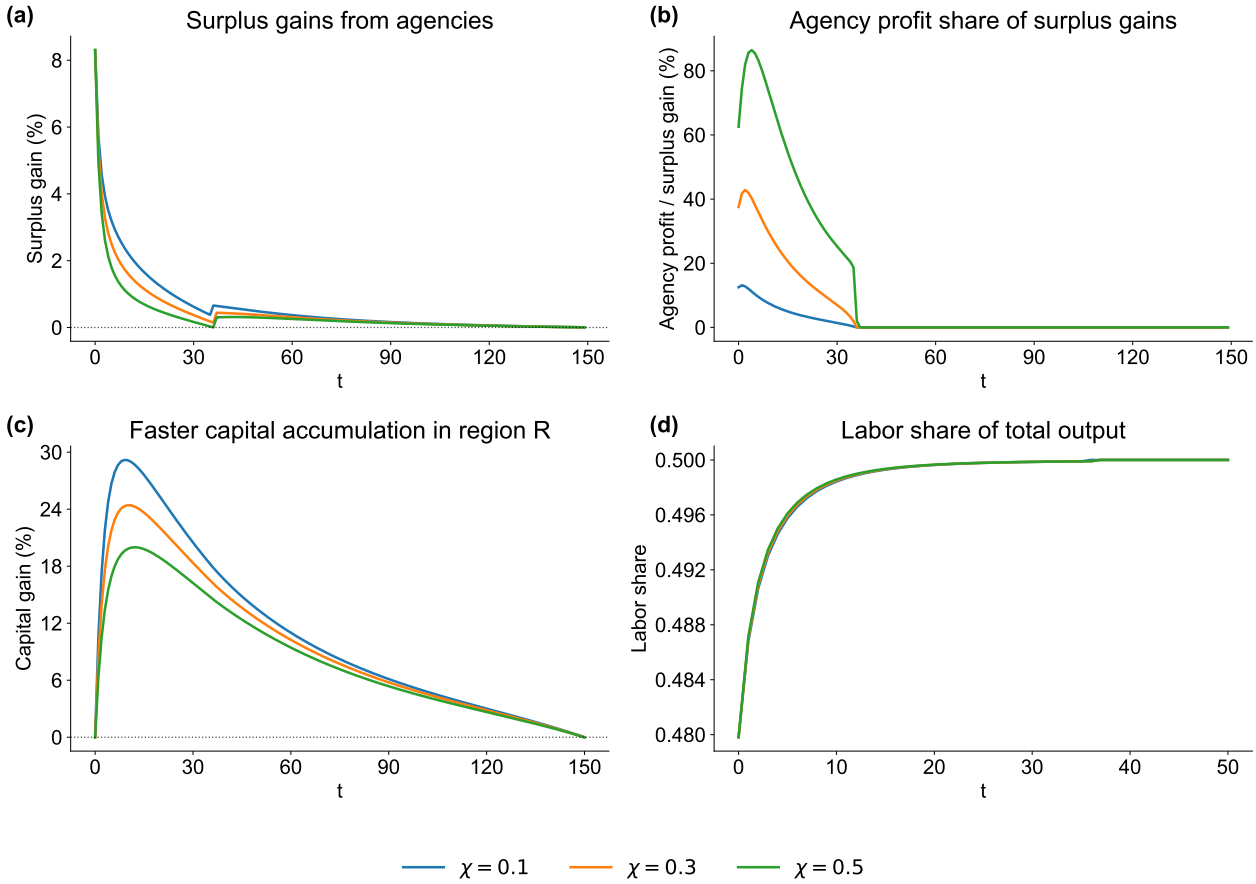
### 4.4.1 Agency Competition in Receiving Region $\chi$

We begin with how different degrees of competition among labor agencies in the labor receiving region affect the path of regional growth and aggregate surplus. Figure 10 shows the transition dynamics in three counterfactual scenarios, $\chi \in \{0.1, 0.3, 0.5\}$.

The variations in $\chi$ do not lead to significant changes in the intensive margins of labor reallocation $L_a$ from both the quantitative results and the theoretical results in Proposition 2. Interestingly however, agency competition in the labor receiving region does affect the aggregate surplus gains. In panel (a) of Figure 10, I plot the relative aggregate surplus gains for different levels of $\chi$, and find that the lower bargaining power of agencies leads to a higher path of aggregate gains. The core insight is that when $\chi = 0.1$, entrepreneurs in the labor receiving regions obtain higher share of the surplus from labor intermediation, which then speeds up capital accumulation and hence generates larger aggregate gains along the convergence path. I plot the dynamics of capital accumulation in region $R$ relative to the no agency counterfactual in panel (c), where the high competition case ($\chi = 0.1$) always implies a faster capital accumulation than the case where agencies have higher bargaining power ($\chi = 0.5$).

Competition among agencies also affect how the surplus gains are distributed. The panel (b) of Figure 10 shows the share of agency profits in aggregate surplus gains. When $\chi = 0.5$, agencies can charge higher price $p_a$ and can capture around 80% of the surplus gains at the initial period, but such profit shares decline as regional wage gap converges and arbitrage opportunities shrink. By comparison, $\chi = 0.1$ yields an overall lower share of aggregate gains being captured by agencies as profits: for example, it brings down such profit share in the initial period from around 80% to

Figure 10: Agency Competition in Labor Receiving Region




*Notes:* The figure shows the dynamics of surplus gains (a), agency profit share of surplus gains (b), capital accumulation (c), and labor output share (d) for $\chi = 0.5$ (green), $\chi = 0.3$ (orange) and $\chi = 0.1$ (blue). In panel (c), capital gain in region $R$ is calculated relative to the no agency counterfactual, i.e., $(K_{Rt} - K_{Rt}^{noagency})/K_{Rt}^{noagency}$. The aggregate surplus with agencies is calculated as $\sum_r y_r - \frac{1}{2}\phi L_a^2 - F$, and the labor income's share of total output is defined as $(\sum_r \sum_s w_r^s L_{sr} + w_a L_a)/\sum_r y_r$.

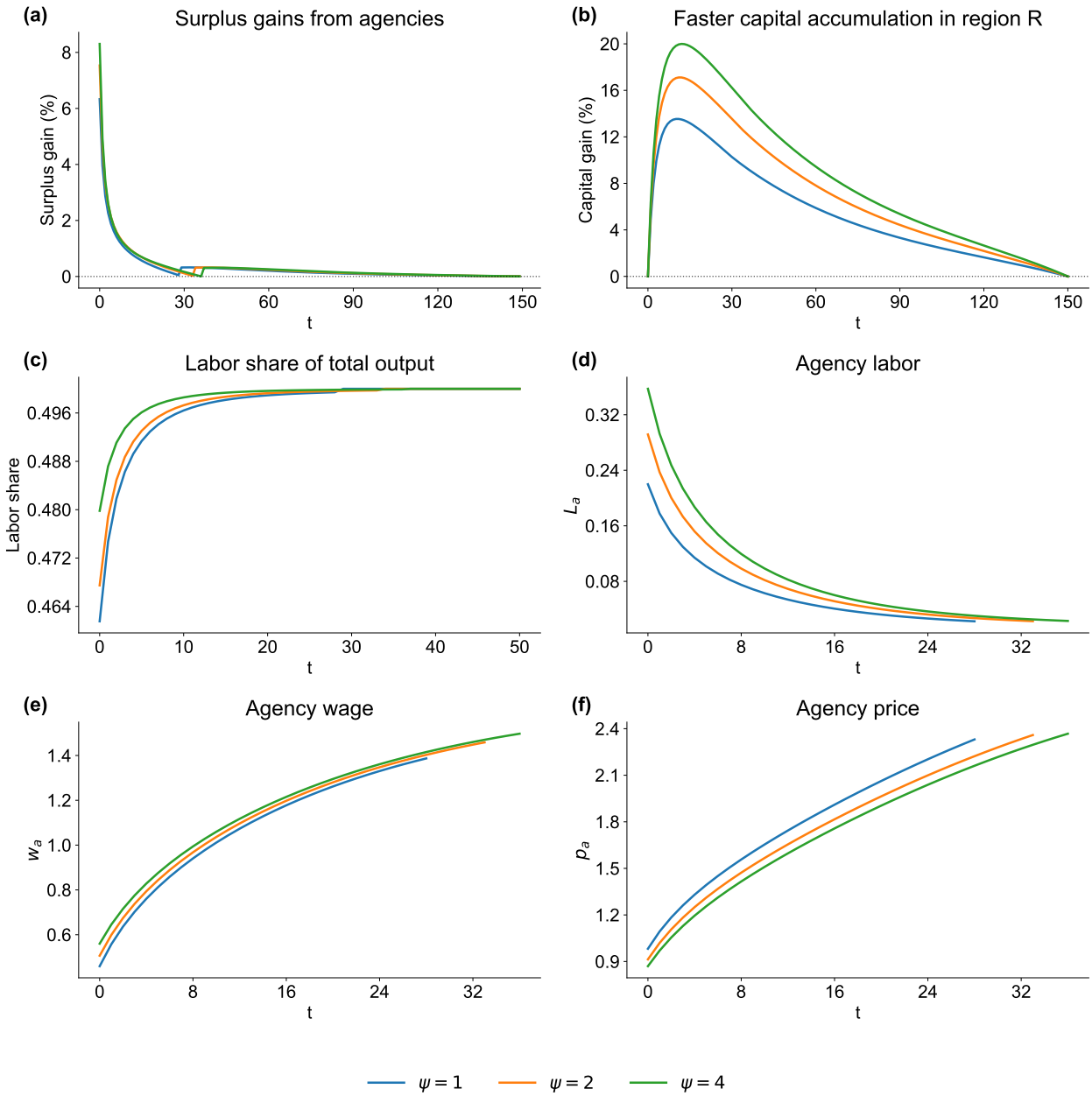
20%. On the other hand, increasing agency competition in the labor receiving region does not lead to significant benefits for workers. In panel (d), I find that labor income share of aggregate output barely changes with different levels of $\chi$.

### 4.4.2 Labor Market Power $\psi$

Figure 11 explores how agencies' labor market power $\psi$ when they hire workers from region $P$ shapes the aggregate gains. In panel (a), lower agency monopsony power increases the aggregate surplus gains in the initial period from 6% to 8%. It also implies a longer periods for agencies to be active. But the mechanism for such surplus gains is different from what we see with $\chi$: while $\chi$ promote surplus gains predominantly from faster capital accumulation, lower labor market power of agencies implies higher incentives to reallocate labor across regions, i.e., higher $L_a$ (panel d). Increasing $\psi = 1$ to $\psi = 4$ can raise the number of unskilled workers hired by agencies from around 0.2 to 0.35. At the same time, there is a lower price that entrepreneurs in region $R$ pay to the agency (panel f), which also tends to speed up capital accumulation (panel b). But the magnitude of speeding up capital accumulation via lower $\chi$ is smaller than what we see from increasing agency competition in region $R$ ($\chi$).

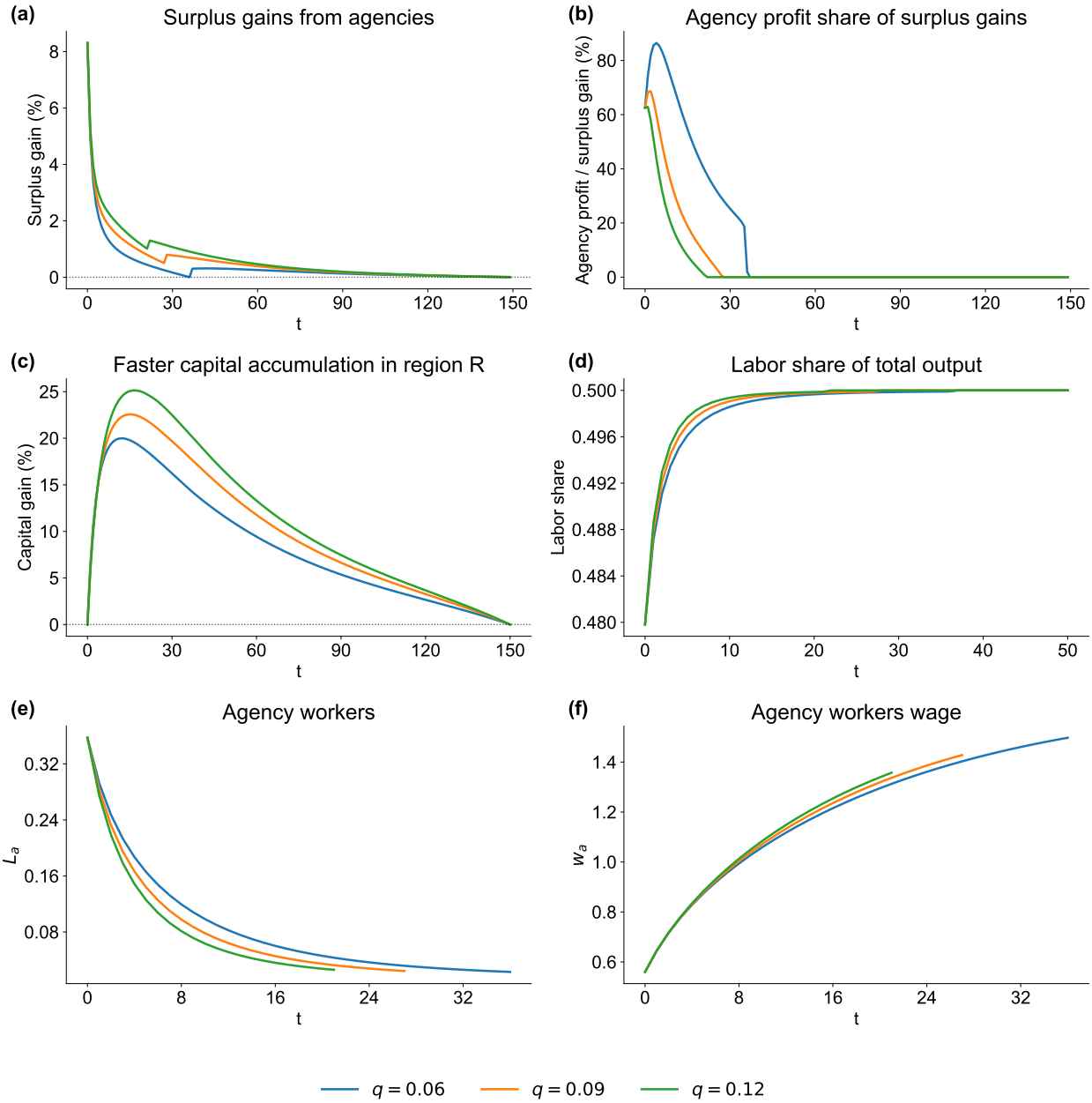
Lower agency monopsony power in region $P$ also implies higher agency worker wage (panel e) and therefore higher labor share of the aggregate output (panel c). When $\psi = 1$, the labor output share at the initial periods is 0.46, lower than 0.48 when $\psi = 4$. This is again in contrast with increasing

Figure 11: Labor Market Power in Labor Sending Region




*Notes:* The figure shows the dynamics of surplus gains (a), capital accumulation (b), labor income share of aggregate output (c), as well as the agency contracts (i.e., $p_a$, $w_a$ and $L_a$) for $\psi \in \{1, 2, 4\}$. The aggregate surplus with agencies is calculated as $\sum_r y_r - \frac{1}{2}\phi L_a^2 - F$, and the labor income's share of total output is defined as $(\sum_r \sum_s w_r^s L_{sr} + w_a L_a)/\sum_r y_r$.

FIGURE 12: AGENCY WORKER TRANSITION




*Notes:* The figure shows the dynamics of surplus gains, agency profit share of the surplus gains, capital accumulation, labor income share of aggregate output, as well as agency contract ($L_a$ and $w_a$) for $q \in \{0.06, 0.09, 0.12\}$. The aggregate surplus with agencies is calculated as $\sum_r y_r - \frac{1}{2}\phi L_a^2 - F$, and the labor income's share of total output is defined as $(\sum_r \sum_s w_r^s L_{sr} + w_a L_a)/\sum_r y_r$.

agency competition in the labor receiving region, which has minimal impact on agency worker and their share of the economic gains.

#### 4.4.3 Agency Workers Transition $q$

We now explore how the probability for agency workers to transition into the permanent labor force in region $R$ affects the regional convergence path. As in panel (a) of Figure 12, increasing such successful transition probability from 6% to 12% can generate substantial aggregate surplus gains along the entire transition path. With higher $q$, more of the agency workers who temporarily stay in region $R$ joins the region as a permanent resident, which speeds up labor transition across regions. In other words, higher $q$ implies that the labor reallocation promoted by agencies is more persistent, i.e., some of the agency workers will successfully stay in the rich region, without having to return to region $P$ and become subject to the mobility barriers again when agencies reduce their employment or exit. In the meantime, agencies exit at an earlier date because of quicker shrinkage of the inter-regional wage disparities.

Panel (e) indicates that since more of the agency workers in earlier periods successfully join the permanent labor force of the rich region, there is lower labor reallocation via agencies in later periods with higher $q$. At the same time, agency workers are paid higher wages (panel f). Because labor transition more quickly from the poor to rich regions, a higher $q$ also leads to faster capital accumulation (panel c).

A higher agency worker transition rate $q$ also helps ensure that the aggregate gains from labor agencies are more widely shared. In panel (b) of Figure 12, agencies can capture at most 80% of the surplus gains along the transition path when $q = 0.06$, in contrast with the fact that agency profit share declines from 60% with $q = 0.12$. Since agency workers are paid higher wages, the total labor income share of aggregate output is also higher along the transition path with $q = 0.12$ than $q = 0.06$.

Taken together, these counterfactual simulations reveal how different structures of the agency sector affect the transition dynamics of regional convergence. Increasing the transition probability of agency workers to permanent workers in the rich region implies that labor reallocations via agencies become more persistent, which then speed up frictional labor transition across regions. Increasing agency competition in labor receiving and sending regions can all generate substantial aggregate surplus gains, but with different underlying mechanisms. Lower agency bargaining power in the labor receiving region implies a larger share of the surplus from labor intermediation accrues to local entrepreneurs, speeding up capital accumulation; on the other hand, lowering agency monopsony power when they hire workers in the poor regions provide stronger incentives for agencies to reallocate more workers, leading to direct effects on surplus gains. Moreover, these policy counterfactual leads to distinctive distributional implications, where higher $q$ and $\psi$ can promote higher labor income share of aggregate output.

## 5 Conclusion

By looking at the empirical context of labor recruitment agencies in China, this paper documents the spatial patterns of agencies expansion, as well as its implications on factor reallocation and regional wage convergence. It then quantifies the aggregate output gains brought by labor agencies using some back-of-the-envelope calculations, and find that around 40% of such aggregate gains are captured by agencies as profits from labor intermediation. Then with a calibrated neoclassical model of regional growth, we see how labor agencies may initially mitigate the inter-regional wage disparities and

speed up the convergence process to the efficient steady state. The quantitative results demonstrate the different distributional implications of increasing agency competition in the labor sending and receiving regions: the former directly incentivizes agencies to reallocate more workers across space; and the latter implies a larger share of surplus accruing to entrepreneurs, speeding up capital accumulation and bring surplus gains. The results also highlight the importance of facilitating the transition of agency worker into the permanent labor force in the receiving regions, which speeds up the Lewis process and increases labor share of aggregate income.

This paper therefore provides a vivid example where resource misallocations may also become economic *opportunities*. It shows that market intermediaries may arise endogenously to arbitrage disparities in prices and marginal product of factor inputs to make profits. In turn this will at least partially correct aggregate misallocations and bring gains in social surplus. It also discusses the extent to which aggregate efficiency gains can be amplified along the dynamic path of economic transition. More importantly, it attempts to provide a new angle to revisit the wedge-based approach to quantify aggregate misallocations, one of the core themes in the macro-development literature. This paper attempts to show the *dynamic* nature of the distortionary wedges, which are often modeled *as if* they are driven by idiosyncratic policies in static settings.

Studying institutions such as these labor recruitment agencies in China opens up a very promising research agenda. For example, in ongoing work, I show that the increasing prevalence of labor agencies in China has enlarged pre-existing regional gaps in educational investment among the youth. This is consistent with the idea that providing job opportunities in richer areas to young people from poorer areas increases their opportunity cost of pursuing higher education, a theme explored as in Atkin (2016) and Rozelle and Hell (2020). This also implies that agencies does not only affect the transition dynamics of capital and labor input, but also shapes human capital accumulation which then have richer implications on aggregate welfare and long-run outcomes. This requires incorporating human capital investment into the dynamic framework to evaluate the welfare impacts of factor reallocations via agencies.

## References


Amodio, F., Medina, P., and Morlacco, M. (2025). Labor market power, self-employment, and development. *American Economic Review*, 115(9):3014–3057.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118.

Atkin, D. (2016). Endogenous skill acquisition and export manufacturing in mexico. *American Economic Review*, 106(8):2046–85.

Atkin, D., Bernadac, B., Donaldson, D., Garg, T., and Huneeus, F. (2024). The incidence of distortions. Working paper.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics*, 21(1):1–266.

Banerjee, A. V. and Duflo, E. (2014). Do firms want to borrow more? testing credit constraints using a directed lending program. *Review of Economic Studies*, 81:572–607.

Banerjee, A. V. and Moll, B. (2010). Why does misallocation persist? *American Economic Journal: Macroeconomics*, 2(1):189–206.

Banerjee, A. V. and Newman, A. F. (1998). Information, the dual economy, and development. *Review of Economic Studies*, 65(4):631–653.

Baqaee, D. R. and Farhi, E. (2019). Productivity and misallocation in general equilibrium. *Quarterly Journal of Economics*, 135(1):105–163.

Barro, R. J., Mankiw, N. G., and Sala-I-Martin, X. (1995). Capital mobility in neoclassical models of growth. *American Economic Review*, 85(1):103–115.

Baseler, T. (2023). Hidden income and the perceived returns to migration. *American Economic Journal: Applied Economics*, 15(4):321–352.

Bau, N. and Matray, A. (2023). Misallocation and capital market integration: Evidence from india. *Econometrica*, 91(1):67–106.

Becker, A.-K. and Hornung, E. (2025). Industrialization and the return to labor: Evidence from prussia. Working paper.

Bertrand, M., Hsieh, C.-T., and Tsivanidis, N. (2024). Contract labor and firm growth in india. Working paper.

Bilal, A. and Lhuillier, H. (2025). Outsourcing, inequality, and aggregate output. Working paper.

Brandt, L., Kambourov, G., and Storesletten, K. (2025). Barriers to entry and regional economic growth in china. *Review of Economic Studies*, pages 1–41.

Bryan, G. and Morten, M. (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy*, 127(5):2229–2268.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225:200–230.

Chen, S., Oliva, P., and Zhang, P. (2022). The effect of air pollution on migration: Evidence from china. *Journal of Development Economics*, 156:102833.

China Labor Watch (2016). Labor rights violations in walmart and home depot's supplier factory. https://chinalaborwatch.org/labor-rights-violations-in-walmart-and-home-depots-supplier-factory/. Accessed: 2025-12-01.

Davila, E. and Schaab, A. (2024). Welfare accounting. Working paper.

de Chaisemartin, C. and D'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*. Forthcoming.

Donaldson, D. and Hornbeck, R. (2016). Railroads and american economic growth: A "market access" approach. *Quarterly Journal of Economics*, 131(2):799–858.

Ebenstein, A. and Zhao, Y. (2015). Tracking rural-to-urban migration in china: Lessons from the 2005 inter-census population survey. *Population Studies*, 69(3):337–353.

Estefan, A., Gerhard, R., Kaboski, J. P., Kondo, I. O., and Qian, W. (2024). Outsourcing policy and worker outcomes. Working paper.

Fang, L. and Herrendorf, B. (2021). High-skilled services and development in china. *Journal of Development Economics*, 151:102671.

Fang, M. and Huang, Z. (2022). Migration, housing constraints, and inequality: A quantitative analysis of china. *Labour Economics*, 78:102200.

Felix, M. and Wong, M. B. (2024). The reallocation effects of domestic outsourcing. Working paper.

Fernando, A. N. and Singh, N. (2025). Regulation by reputation? intermediaries, labor abuses, and international migration. *Review of Economics and Statistics*. Forthcoming.

Gai, Q., Guo, N., Li, B., Shi, Q., and Zhu, X. (2025). Rural pensions, labor reallocation, and aggregate income: An empirical and quantitative analysis of china. *Econometrica*, 93(5):1663–1696.

Garicano, L., Lelarge, C., and Reenen, J. V. (2016). Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review*, 106(11):3439–79.

Gennaioli, N., Porta, R. L., Lopez-De-Silanes, F., and Shleifer, A. (2013). Human capital and regional development. *Quarterly Journal of Economics*, 128(1):105–164.

Gennaioli, N., Porta, R. L., Lopez-De-Silanes, F., and Shleifer, A. (2014). Growth in regions. *Journal of Economic Growth*, 19:259–309.

Ghatak, M. (2015). Theories of poverty traps and anti-poverty policies. *World Bank Economic Review*, 29:s77–s105.

Ghatak, M. and Mookherjee, D. (2024). Misallocating misallocation? Working paper.

Gong, X., Qiao, Z., and Wright, R. (2025). Middlemen in search equilibrium: A survey. Working paper.

Harris, H. J. (1982). *The Right to Manage: Industrial Relations Policies of American Business in the 1940s*. University of Wisconsin Press.

Hatton, E. (2011). *The Temp Economy: From Kelly Girls to Permatemps in Postwar America*. Temple University Press.

Heise, S. and Porzio, T. (2023). Labor misallocations across firms and regions. Working paper.

Holmes, T. J. and Snider, J. T. (2011). A theory of outsourcing and wage decline. *American Economic Journal: Microeconomics*, 3:38–59.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Huang, P. C. C. (2017). Dispatch work in china: A study from case records, part i. *Modern China*, 43(3):247–287.

Itskhoki, O. and Moll, B. (2019). Optimal development policies with financial frictions. *Econometrica*, 87(1):139–173.

Kalleberg, A. L. (2011). *Good Jobs, Bad Jobs: The Rise of Polarized and Precarious Employment Systems in the United States, 1970s-2000s*. Russell Sage Foundation.

Kleinman, B., Liu, E., and Redding, S. J. (2023). Dynamic spatial general equilibrium. *Econometrica*, 91(2):385–424.

Kline, P. (2025). Labor market monopsony: Fundamentals and frontiers. In Dustmann, C. and Lemieux, T., editors, *Handbook of Labor Economics*, pages 655–728. Elsevier.

Lagakos, D. (2020). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic Perspectives*, 34(3):174–192.

Lagakos, D., Mobarak, A. M., and Waugh, M. E. (2023). The welfare effects of encouraging rural-urban migration. *Econometrica*, 91(3):803–837.

Lau, L. J., Qian, Y., and Roland, G. (2000). Reform without losers: An interpretation of china's dual-track approach to transition. *Journal of Political Economy*, 108(1):120–143.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, 22(2):139–191.

Lin, J. Y. (1992). Rural reforms and agricultural growth in china. *American Economic Review*, 82(1):34–51.

Liu, G. (2014). Private employment agencies and labor dispatch in china. ILO Working paper.

Lucas, R. E. (1990). Why doesn't capital flow from rich to poor countries. *American Economic Review: Papers and Proceedings*, 80(2):92–93.

Meng, X. (2012). Labor market outcomes and reforms in china. *Journal of Economic Perspectives*, 26(4):75–102.

Micco, A. and Munoz, P. (2024). The impact of extending employment protection to agency workers on firms. *American Economic Journal: Macroeconomics*, 16(1):66–101.

Moll, B. (2014). Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review*, 104(10):3186–3221.

Munoz, M. (2024). Trading non-tradables: The implications of europe's job-posting policy. *Quarterly Journal of Economics*, 139(1):235–304.

Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration and the rural-urban wage gap. *American Economic Review*, 106(1):46–98.

Ngai, L. R., Pissarides, C. A., and Wang, J. (2019). China's mobility barriers and employment allocations. *Journal of the European Economic Association*, 17(5):1617–1653.

Rankin, N., Soderbom, M., and Teal, F. (2006). Exporting from manufacturing firms in sub-saharan africa. *Journal of African Economies*, 15(4):671–687.

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11:707–720.

Rozelle, S. and Hell, N. (2020). *Invisible China: How the Urban-Rural Divide Threatens China's Rise*. The University of Chicago Press.

Schumpeter, J. A. (1934). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Harvard University Press, Cambridge, MA.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1):65–94.

Sraer, D. and Thesmar, D. (2023). How to use natural experiments to estimate misallocation. *American Economic Review*, 113(4):906–938.

State Council (2006). Several opinions of the state council on issues concerning migrant workers. Technical report, State Council of People's Republic of China, Beijing, China.

Tombe, T. and Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–1872.

Tu, W. (2017). The changing role of government towards labour. In Brown, W. and Kai, C., editors, *The Emerging Industrial Relations in China*, pages 112–140. Cambridge University Press, Cambridge.

Wang, F. (2019). *Labor Dispatch in China: Three Essays*. PhD thesis, Rutgers, The State University of New Jersey.

Yan, W. (2023). Evolution and the impetus mechanisms of the labour dispatch in china. *World Economy*, 11:128–154.

Young, A. (2000). The razor's edge: Distortions and incremental reform in the people's republic of china. *Quarterly Journal of Economics*, 115(4):1091–1135.

Young, A. (2013). Inequality, the urban-rural gap, and migration. *Quarterly Journal of Economics*, 128(4):1727–1785.

# Appendix

# A Empirical Results

## A.1 Summary Statistics

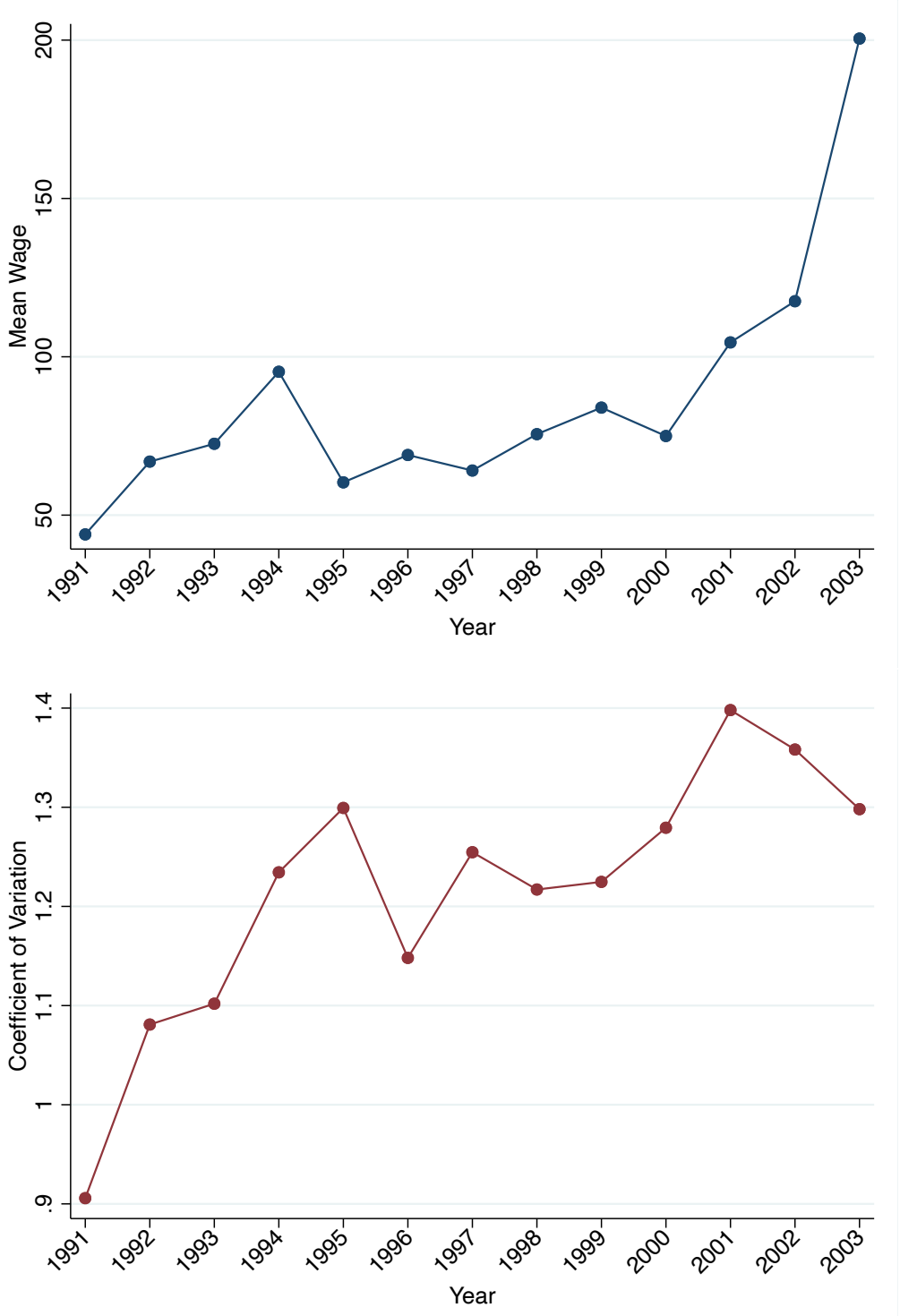
Table A.1: Summary Statistics

| Variable | Mean | Std. Dev. |
|---|---|---|
| **Data Source: Administrative Tax Survey** | | |
| City average ex ante wage (manufacturing) | 17.02 | 0.22 |
| City average wage (manufacturing) | 28.26 | 0.17 |
| City average MPL (manufacturing) | 24.86 | 0.13 |
| City average capital-labor ratio (manufacturing) | 136.06 | 1.18 |
| City average ex ante wage (comparable sectors) | 21.88 | 0.51 |
| City average wage (comparable sectors) | 34.65 | 0.19 |
| Agency headquarter employment | 864.00 | 1846.41 |
| Agency headquarter profits | 23483.24 | 76956.10 |
| **Data Source: China Labor Dynamics Survey** | | |
| City average annual wage (college) | 52.04 | 29.81 |
| City average annual wage (below college) | 25.99 | 12.44 |
| **Data Source: Chinese City Statistical Yearbook** | | |
| City ex ante (2005) wage | 15.51 | 4.71 |
| City ex ante (2007) wage | 21.22 | 5.80 |
| City ex ante (2005) output share | 0.35% | 0.005 |
| Fixed assets of industrial sector (2005) | 28816.11 | 46877.13 |
| City wage (2008-2018) | 46.21 | 18.91 |
| *Hukou* population (2008-2018) | 429.04 | 305.29 |
| Number of High School (2008-2018) | 214.33 | 137.57 |
| High School Student-Teacher Ratio (2008-2018) | 13.49 | 3.32 |

*Notes:* All wages and agency profits are measured in thousands nominal RMB. To obtain manufacturing city average MPL, I first calculate $\beta\frac{Y}{L}$ for each firm where $\beta$ is the two-digit industry labor share, $Y$ is the total output and $L$ is the total employment, and I then take city average (weighted by firm employment). City average capital-labor ratio is the ratio of total fixed assets of manufacturing firms over total employment. *Hukou* population is measured in 10 thousands, and fixed assets of industrial sector (2005) is measured in 10 thousands RMB.

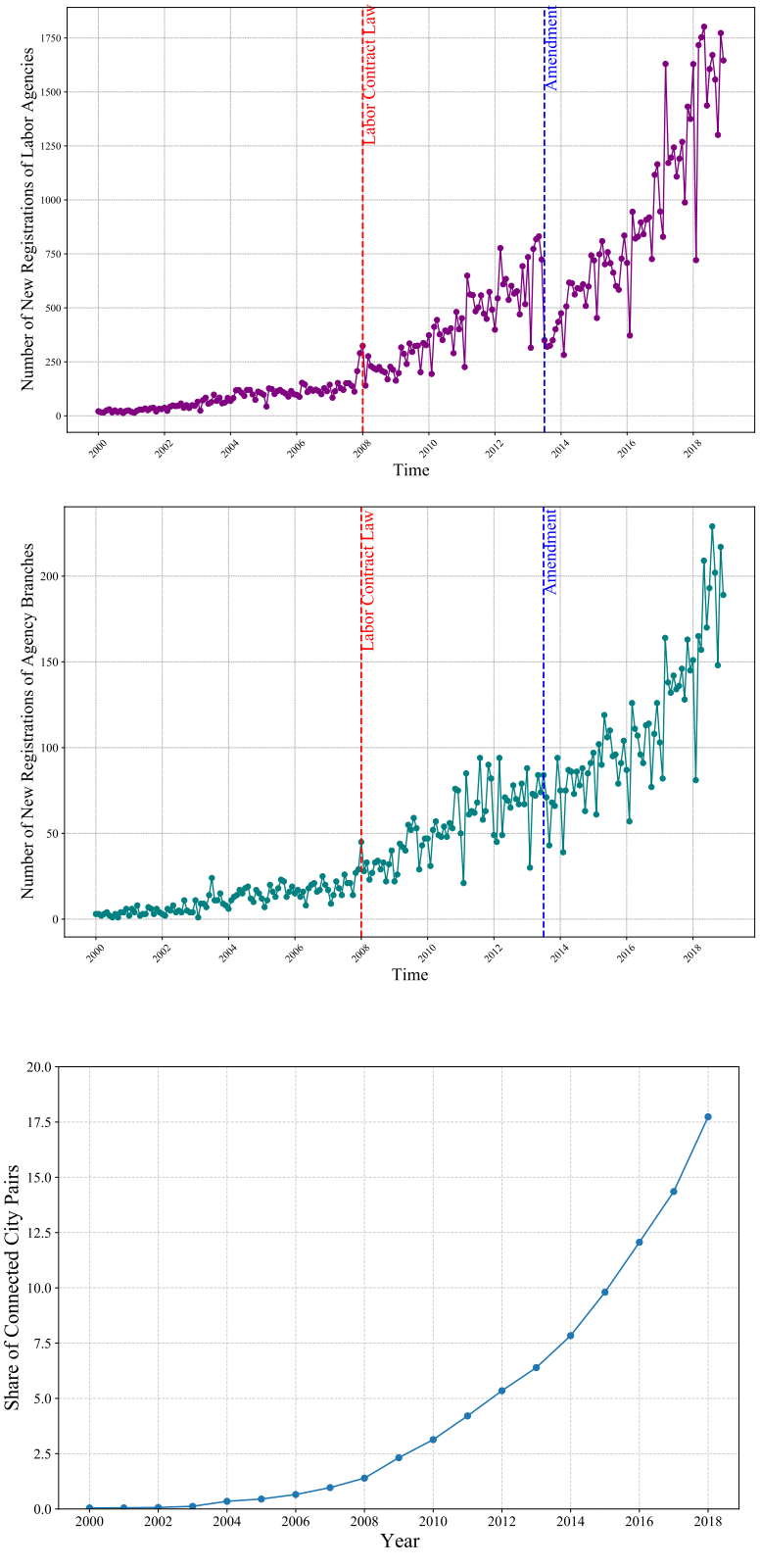
## A.2 Descriptive Facts

Figure A.1: Wage Growth and Wage Dispersion in Sub-Saharan Africa



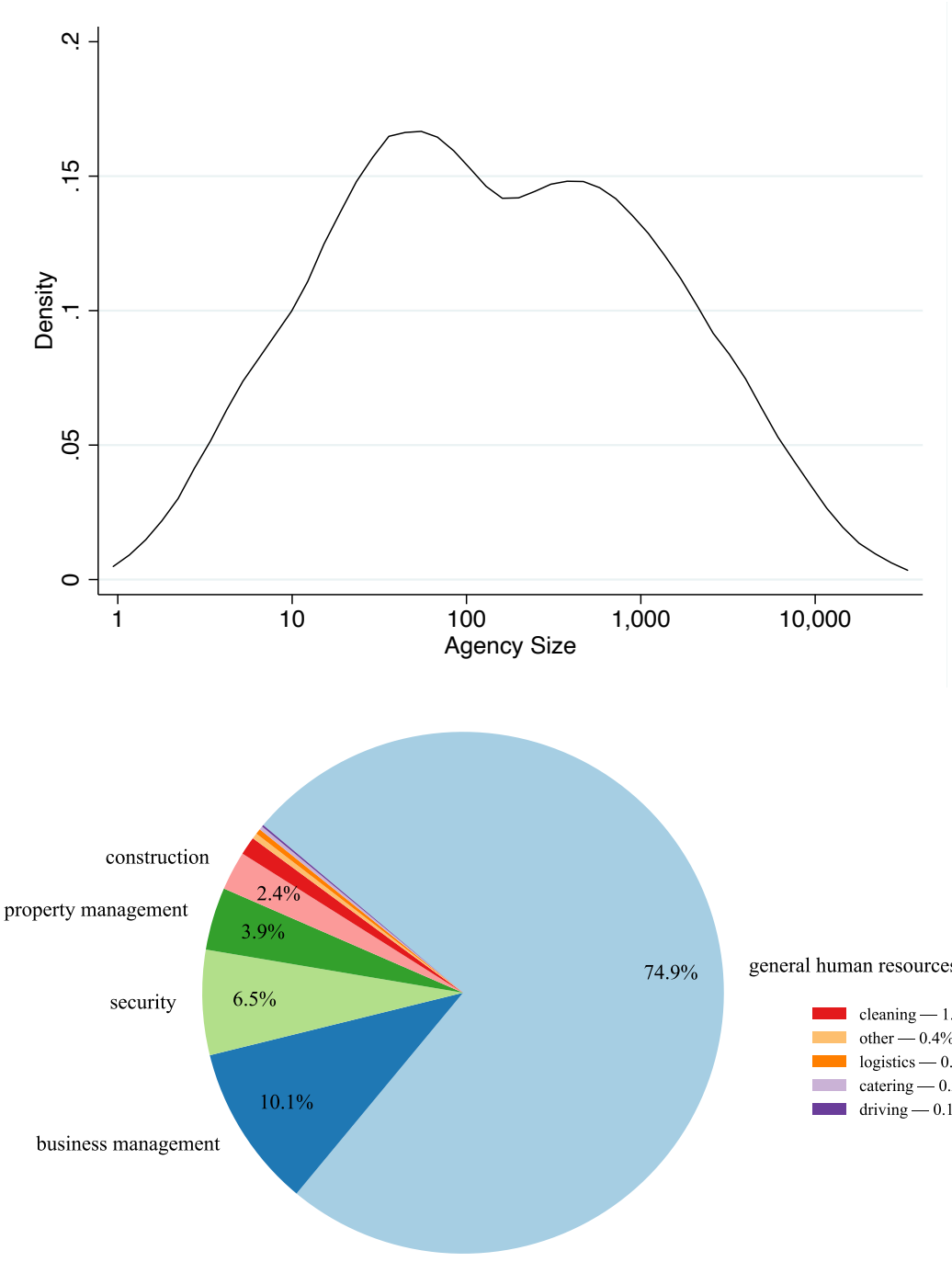

*Notes:* The figure shows average firm wage (top panel) and the coefficient of variation of firm wage (bottom panel) between 1991 and 2003, based on the sample of manufacturing firms in Ghana, Kenya, and Nigeria in Rankin et al. (2006).

FIGURE A.2: LABOR AGENCIES: ENTRY OVER TIME




*Notes:* The top (middle) panel shows the monthly time series of registrations of labor recruitment agencies (agency branches or subsidiaries that belong to inter-regional agency groups) in China over 2000-2019. The red dashed line captures the timing of the introduction of Labor Contract Law in January 2008, and the blue dashed line captures the Amendment to Labor Contract Law in July 2013. The bottom panel shows the share of city pairs that are connected by agency groups over time.

Figure A.3: Labor Agencies: Size and Specialization




*Notes:* The top panel is the density plot of employment of agency group headquarters, based on the sub-sample of agencies in the tax survey dataset. The bottom panel presents the categorization of agency groups and their local branches, based on their business range information on *aiqicha* and the author's own classification.

## A.3 Agencies Expansion

Table A.2: ex ante Wage and Agency Expansion

| | (1) | (2) | (3) |
|---|---|---|---|
| | Branched | Branched | Branched |
| Bottom 10% | 0.00372*** | 0.00377*** | 0.195*** |
| | (0.000270) | (0.000269) | (0.0582) |
| Top 10% | 0.0112*** | 0.0113*** | 0.288*** |
| | (0.000576) | (0.000577) | (0.0420) |
| Distance | -0.0447*** | -0.0452*** | -1.437*** |
| | (0.000382) | (0.000376) | (0.0245) |
| Branched Before 2007 | 0.712*** | | |
| | (0.00561) | | |
| High School Student Teacher Ratio | -0.000465*** | -0.000478*** | 0.0628*** |
| | (0.0000395) | (0.0000395) | (0.00657) |
| Number of High Schools | -0.00583*** | -0.00586*** | -0.0798** |
| | (0.000206) | (0.000206) | (0.0321) |
| Library Collections | 0.00683*** | 0.00684*** | 0.647*** |
| | (0.000209) | (0.000208) | (0.0211) |
| Observations | 906449 | 905467 | 888023 |
| Methods | OLS | OLS | Conditional Logit |
| *Fixed Effects* | | | |
| Agency Group | Yes | Yes | Yes |

*Notes:* This table shows the regression results for the relationship between ex ante city average wage and agency expansion decision, where Bottom 10% and Top 10% are binary indicator for whether the ex ante (i.e., 2007) city average wage in the Chinese City Statistical Yearbook belongs to the bottom 10% and top 10% decile group. Branched is a binary variable which equals 1 if an agency group has ever established a branch in the city during 2008-2018. Branched Before 2007 is a binary variable which equals to 1 if the agency group had branch in the city before 2007. The first and the second column is estimated using the linear probability model, and the third column is obtained from a conditional logit model. The second and the third columns are based on observations that don't have branch in the city before 2007. Standard errors reported in parentheses are clustered at agency group level: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

## A.4 Comparable Sectors

Table A.3: Agency Network and Regional Wage Gap (Comparable Sectors): 2008 - 2016

| | (1) Construction | (2) Construction | (3) Hospitality | (4) Hospitality | (5) Logistics& Warehousing | (6) Logistics& Warehousing |
|---|---|---|---|---|---|---|
| Connected | 0.0399*** | 0.0370*** | 0.0354*** | 0.0261*** | 0.0322*** | 0.0261*** |
| | (0.00617) | (0.00648) | (0.00813) | (0.00846) | (0.00818) | (0.00834) |
| Connected×$\lvert\Delta Wage^{pre}\rvert$ | -0.100*** | -0.0816*** | -0.138*** | -0.111*** | -0.134*** | -0.112*** |
| | (0.0144) | (0.0154) | (0.0242) | (0.0252) | (0.0182) | (0.0188) |
| Connections | | 0.00166 | | 0.00539*** | | 0.00353** |
| | | (0.00122) | | (0.00172) | | (0.00157) |
| Connections×$\lvert\Delta Wage^{pre}\rvert$ | | -0.0139*** | | -0.0161*** | | -0.0148*** |
| | | (0.00368) | | (0.00427) | | (0.00416) |
| Observations | 66765 | 66765 | 10965 | 10965 | 28841 | 28841 |
| *Fixed Effects* | | | | | | |
| City Pair | Yes | Yes | Yes | Yes | Yes | Yes |
| Year | Yes | Yes | Yes | Yes | Yes | Yes |

**Notes:** This table reports the OLS estimates of the effect of LDA expansion on wage disparities across cities in correspondingly construction, hospitality, and logistics/warehousing sectors during 2008-2016. The results are obtained based on the administrative dataset of tax survey of firms. Outcome variables are the absolute value of difference of (log) average wage for the corresponding sector between two cities. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. Connection is the number of such agency groups. $\lvert\Delta Wage^{pre}\rvert$ is the absolute value of ex ante (i.e., 2007) log average wage difference for the corresponding sector. Standard errors reported in parentheses are clustered at city pair level: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

## A.5 Factor Inputs

Table A.4: Agency Network and Regional MPL Gap: 2008 - 2016

| | (1) $\lvert\Delta MPL\rvert$ | (2) $\lvert\Delta MPL\rvert$ | (3) $\lvert\Delta MPL\rvert$ |
|---|---|---|---|
| Connected | 0.00815*** | 0.0482*** | 0.0430*** |
| | (0.00193) | (0.00295) | (0.00328) |
| Connected×$\lvert\Delta MPL^{pre}\rvert$ | | -0.158*** | -0.127*** |
| | | (0.00956) | (0.0115) |
| Connection | | | 0.00362*** |
| | | | (0.00106) |
| Connection×$\lvert\Delta MPL^{pre}\rvert$ | | | -0.0230*** |
| | | | (0.00462) |
| Observations | 171250 | 171250 | 171250 |
| *Fixed Effects* | | | |
| City Pair | Yes | Yes | Yes |
| Year | Yes | Yes | Yes |

**Notes:** This table reports the OLS estimates of the effect of agency network expansion on manufacturing MPL disparities across cities during 2008-2016. The results are obtained based on the administrative dataset of tax survey of firms. $\lvert\Delta MPL\rvert$ is the absolute value of difference of log average (weighted by employment) MPL between two cities. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. Connection is the number of such agency groups. $\lvert\Delta MPL^{pre}\rvert$ is the absolute value of ex ante (i.e., 2007) log average MPL difference. Standard errors reported in parentheses are clustered at city pair level: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Table A.5: Agency Network and Regional Capital-Labor Ratio Gap: 2008 - 2016

| | (1) $\Delta(K/L)$ | (2) $\Delta(K/L)$ | (3) $\Delta(K/L)$ |
|---|---|---|---|
| Connected | -0.00538 | -0.00475 | -0.00364 |
| | (0.00448) | (0.00446) | (0.00479) |
| Connected×$\Delta(K/L)^{pre}$ | | -0.0698*** | -0.0666*** |
| | | (0.0105) | (0.0109) |
| Connection | | | -0.000874 |
| | | | (0.00150) |
| Connection×$\Delta(K/L)^{pre}$ | | | -0.00242 |
| | | | (0.00269) |
| Observations | 171250 | 171250 | 171250 |
| *Fixed Effects* | | | |
| City Pair | Yes | Yes | Yes |
| Year | Yes | Yes | Yes |

**Notes:** This table reports the effect of agencies expansion on (manufacturing) capital-labor ratio gap across cities (2008-2016), based on the tax survey dataset. $\Delta(K/L)$ is the (log) average capital-labor ratio difference. Connected is a binary variable which equals to 1 if any agency group has active branches in both cities. Connection is the number of such agency groups. $\Delta(K/L)^{pre}$ is the ex ante (i.e., 2007) log average capital-labor ratio difference. Standard errors reported in parentheses are clustered at city pair level: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

## A.6 Determinants of $\Theta_{mn}$

Table A.6: $\Theta_{mn}$ and Predetermined City-Pair Characteristics

| | $\Theta_{mn}$ |
|---|---|
| GDP per capita gap | 0.013 |
| | (0.027) |
| Agricultural employment share gap | 0.001 |
| | (0.003) |
| Population gap | -0.005 |
| | (0.019) |
| Log road distance | -0.033 |
| | (0.023) |
| Log flow in 2010 | 0.012 |
| | (0.009) |
| Observations | 2,189 |
| F-statistic | 1.13 |
| *Fixed Effects* | |
| Province Pair | Yes |

*Notes:* This table reports the relationship between the instrument $\Theta_{mn}$ and predetermined city-pair characteristics. The dependent variable is $\Theta_{mn}$, defined as the share of agency connections via local offices for city pair $mn$. All specifications absorb province-pair fixed effects. Standard errors reported in parentheses are clustered at the province-pair level. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

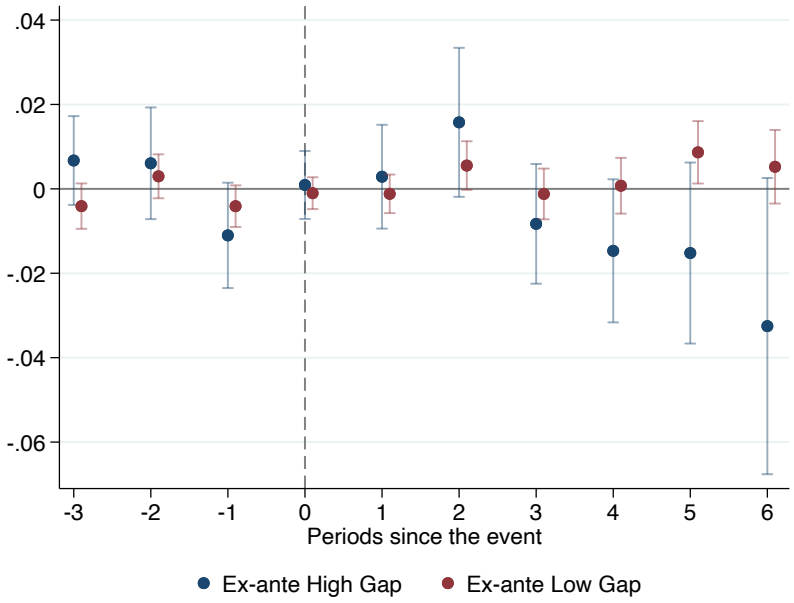
## A.7 Spillovers

Table A.7: Agency Network and Regional Wage Gap: Sub-Sample (I)

| | $\lvert\Delta Wage\rvert$ | $\lvert\Delta Wage\rvert$ | $\lvert\Delta Wage\rvert$ |
|---|---|---|---|
| Connected | -0.0104*** | 0.0445*** | 0.0385*** |
| | (0.00153) | (0.00198) | (0.00263) |
| Connected$\times\lvert\Delta Wage^{pre}\rvert$ | | -0.278*** | -0.244*** |
| | | (0.00893) | (0.0145) |
| Connections | | | 0.00399*** |
| | | | (0.00116) |
| Connections$\times\lvert\Delta Wage^{pre}\rvert$ | | | -0.0245*** |
| | | | (0.00815) |
| Observations | 146922 | 146922 | 146922 |
| *Fixed Effects* | | | |
| City Pair | Yes | Yes | Yes |
| Year | Yes | Yes | Yes |

**Notes:** This table reports the OLS estimates of LDA expansion on manufacturing wage disparity across cities (2008-2016). This result is based on the average wage gap for each city pairs calculated from the tax survey dataset, but I remove the potential geographical spillover control groups.$|\Delta Wage|$ is the absolute value of difference of (log) average wage between two cities. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. Connection is the number of such agency groups. $|\Delta Wage^{pre}|$ is the absolute value of ex ante (i.e., 2007) log wage difference. I exclude control city pairs specified in $S^{G}_{mnt}$. Standard errors reported in parentheses are clustered at city pair level: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***} < 0.01$.

Figure A.4: Agency Network and Regional Wage Gap: Event Study with Sub-Sample (I)




*Notes:* The figure shows estimates from the event study specification of difference of (log) average (manufacturing) wage for city pairs on Connected (i.e., if there exist an agency group that has active branches in both cities), respectively for city pairs with above and below 75th ex ante percentile wage disparity, based on the methodology of Callaway and Sant'Anna (2021). I exclude control city pairs specified in $S^G_{mnt}$. I show the 95% confidence interval calculated using standard error clustered at city pair level.

Table A.9: Agency Network and Regional Wage Gap: Sub-Sample (II)

| | $\lvert\Delta Wage\rvert$ | $\lvert\Delta Wage\rvert$ | $\lvert\Delta Wage\rvert$ |
|---|---|---|---|
| Connected | -0.0148*** | 0.0546*** | 0.0474*** |
| | (0.00155) | (0.00187) | (0.00249) |
| Connected×$\lvert\Delta Wage^{pre}\rvert$ | | -0.329*** | -0.290*** |
| | | (0.00724) | (0.0133) |
| Connections | | | 0.00494*** |
| | | | (0.00111) |
| Connections×$\lvert\Delta Wage^{pre}\rvert$ | | | -0.0277*** |
| | | | (0.00808) |
| Observations | 150790 | 150790 | 150790 |
| *Fixed Effects* | | | |
| City Pair | Yes | Yes | Yes |
| Year | Yes | Yes | Yes |

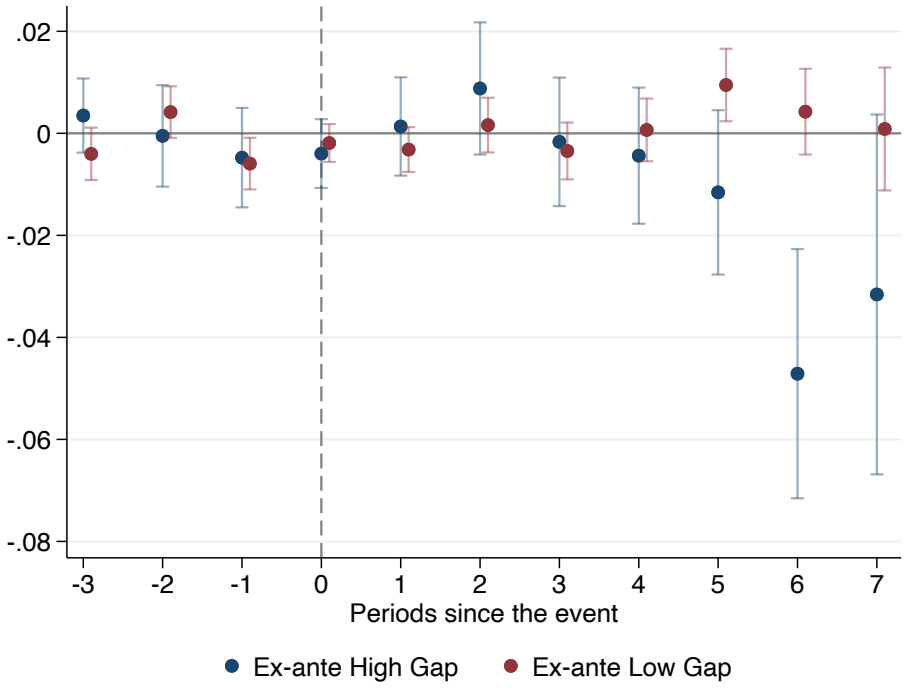
**Notes:** This table reports the OLS estimates of LDA expansion on manufacturing wage disparity across cities (2008-2016). This result is based on the average wage gap for each city pairs calculated from the tax survey dataset, but I remove the potential geographical spillover control groups.$\lvert\Delta Wage\rvert$ is the absolute value of difference of (log) average wage between two cities. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. Connection is the number of such agency groups. $\lvert\Delta Wage^{pre}\rvert$ is the absolute value of ex ante (i.e., 2007) log wage difference. I exclude control city pairs with $\hat{\Psi}_{mnt}$ above the 90th percentile of $\hat{\Psi}_{mnt}$ distribution. Standard errors reported in parentheses are clustered at city pair level: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Table A.8: Agency Network and Labor Mobility Across Cities: Sub-Sample

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Gross Flow | Gross Flow | Gross Flow | Gross Flow (College) | Gross Flow (Below College) |
| Connected | 7.064*** | 7.079*** | 4.593*** | 1.570*** | 2.888*** |
| | (1.040) | (1.030) | (1.322) | (0.371) | (0.989) |
| Connected×$\Delta Wage^{pre}$ | | 19.87*** | 11.01 | 3.148** | 7.475 |
| | | (4.616) | (6.721) | (1.467) | (5.187) |
| Connection | | | 1.744*** | 0.646*** | 1.041** |
| | | | (0.640) | (0.198) | (0.480) |
| Connection×$\Delta Wage^{pre}$ | | | 6.312* | 1.297* | 4.859* |
| | | | (3.366) | (0.732) | (2.636) |
| Observations | 147632 | 146612 | 146612 | 146612 | 146612 |
| *Fixed Effects* | | | | | |
| City Pair | Yes | Yes | Yes | Yes | Yes |
| Year | Yes | Yes | Yes | Yes | Yes |

**Notes:** This table reports the OLS estimates of the effect of expansion of agency network on worker mobility across cities. The results are obtained based on the sub-sample of population census in 2005 and 2010. $\Delta Wage^{pre}$ is the ex ante (i.e., 2005) log average wage difference between city $n$ and $m$. Gross flow from city $m$ to city $n$ (where $n \neq m$) is calculated as the total headcount (in thousands) for which the respondents reported to have lived in city $n$ (for more than one year) but their hukou registration is in city $m$. Gross flow of college (below college) labor is measured by such flow for which the respondents have a college (below college) degree. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. I exclude control city pairs specified in $S^G_{mnt}$. Standard errors reported in parentheses are clustered at city pair level: *$p < 0.10$, *$p < 0.05$, *$p < 0.01$.

Figure A.5: Agency Network and Regional Wage Gap: Event Study with Sub-Sample (II)




*Notes:* The figure shows estimates from the event study specification of difference of (log) average (manufacturing) wage for city pairs on Connected (i.e., if there exist an agency group that has active branches in both cities), respectively for city pairs with above and below 75th ex ante percentile wage disparity, based on the methodology of Callaway and Sant'Anna (2021). I exclude control city pairs with $\hat{\Psi}_{mnt}$ above the 90th percentile of $\hat{\Psi}_{mnt}$ distribution. I show the 95% confidence interval calculated using standard error clustered at city pair level.

## A.8 Average Wage

Table A.10: Agency Network and Regional Wage Gap (Yearbook Data): 2008 - 2016

| | $\lvert\Delta Wage\rvert$ | $\lvert\Delta Wage\rvert$ | $\lvert\Delta Wage\rvert$ |
|---|---|---|---|
| Connected | -0.00398*** | 0.0421*** | 0.0374*** |
| | (0.00119) | (0.00165) | (0.00219) |
| Connected×$\lvert\Delta Wage^{pre}\rvert$ | | -0.123*** | -0.110*** |
| | | (0.00404) | (0.00502) |
| Connections | | | 0.00316*** |
| | | | (0.00102) |
| Connections×$\lvert\Delta Wage^{pre}\rvert$ | | | -0.00841*** |
| | | | (0.00203) |
| Observations | 445012 | 428601 | 428601 |
| *Fixed Effects* | | | |
| City Pair | Yes | Yes | Yes |
| Year | Yes | Yes | Yes |

**Notes:** This table reports the OLS estimates of LDA expansion on average wage disparity across cities (2008-2016), based on the Chinese City Statistical Yearbook. $\lvert\Delta Wage\rvert$ is the absolute value of difference of (log) average wage between two cities. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. Connection is the number of such agency groups. $\lvert\Delta Wage^{pre}\rvert$ is the absolute value of ex ante (i.e., 2007) log average wage difference. Standard errors reported in parentheses are clustered at city pair level: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

## A.9 *Hukou* Population Gap

Table A.11: Agency Network and *Hukou* Population Gap: 2008 - 2016

| | $\Delta Hukou$ | $\Delta Hukou$ | $\Delta Hukou$ |
|---|---|---|---|
| Connected | 0.00389*** | -0.00284** | -0.000549 |
| | (0.00113) | (0.00111) | (0.00159) |
| Connected$\times\Delta Wage^{pre}$ | | 0.0885*** | 0.0676*** |
| | | (0.00279) | (0.00561) |
| Connection | | | -0.00175** |
| | | | (0.000877) |
| Connection$\times\Delta Wage^{pre}$ | | | 0.0141*** |
| | | | (0.00338) |
| Observations | 454959 | 435550 | 435550 |
| *Fixed Effects* | | | |
| City Pair | Yes | Yes | Yes |
| Year | Yes | Yes | Yes |

**Notes:** This table reports the OLS estimates of the effect of LDA expansion on *Hukou* population gap across cities during 2008-2016. The results are obtained based on the Chinese City Statistical Yearbook. $\Delta Hukou$ is the difference of log *Hukou* population between city $m$ and $n$. $\Delta Wage$ is the difference of (log) average wage between city $m$ and city $n$. Connected is a binary variable which equals to 1 if there exist one agency group that has active branches in both cities. Connection is the number of such agency groups. $\Delta Wage^{pre}$ is the ex ante (i.e., 2007) log average wage difference. Standard errors reported in parentheses are clustered at city pair level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## A.10 Differential Trends in Labor Flow

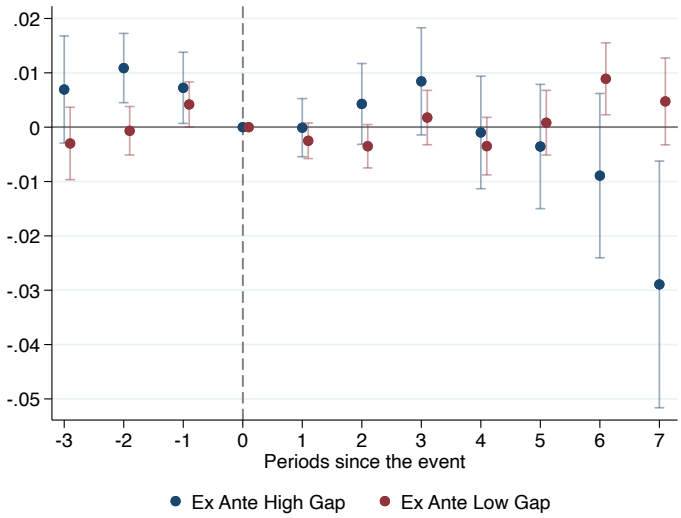
Table A.12: Pre-Existing Differential Trends in Labor Flow: 2000 and 2005

| | Flow | Flow | Flow |
|---|---|---|---|
| Treat$\times\mathbb{1}[t = 2005]$ | 0.333 | 0.166 | 0.316 |
| | (0.351) | (0.658) | (0.411) |
| Treat | 3.414*** | 3.962*** | 3.093*** |
| | (0.175) | (0.376) | (0.185) |
| $\mathbb{1}[t = 2005]$ | 0.436*** | 1.710*** | 0.0203* |
| | (0.0334) | (0.130) | (0.0117) |
| Observations | 133187 | 33187 | 100000 |

**Notes:** This table reports the OLS estimates of equation (7), where $Treat$ is a binary variable which equals to one if city pair $mn$ is connected in 2010, and zero otherwise. $\mathbb{1}[t = 2005]$ is an indicator variable which equals to one if year is 2005. The first column is based on the full sample, and the second (third) column is based on city pairs with above (below) 75th percentile ex ante wage gap. Standard errors reported in parentheses are clustered at city pair level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## A.11 Alternative DiD Estimator

Figure A.6: Event Study: Agency Network and Regional Wage Gap




*Notes:* The figure shows estimates from the event study specification of difference of (log) average (manufacturing) wage for city pairs on Connected (i.e., if there exist an agency group that has active branches in both cities), respectively for city pairs with ex ante above and below 75% wage disparity, based on de Chaisemartin and D'Haultfoeuille (2024).

## A.12 Output Gains

Table A.13: Relative Output Gains: Linear Projections

| | $y = AK^{\alpha}H^{1-\alpha}$ | $y = AK^{\alpha}(H_U)^{\beta}(H_S)^{1-\alpha-\beta}$ | | |
|---|---|---|---|---|
| | | $\beta = 0.2$ | $\beta = 0.4$ | $\beta = 0.6$ |
| $\{\text{Connections}\}_{2011}$ | 8.85% | 5.42% | 7.06% | 8.85% |
| $\{\text{Connections}\}_{2012}$ | 9.97% | 7.30% | 8.79% | 10.45% |
| $\{\text{Connections}\}_{2013}$ | 10.86% | 8.47% | 9.88% | 11.53% |
| $\{\text{Connections}\}_{2014}$ | 12.43% | 9.95% | 11.36% | 13.09% |
| $\{\text{Connections}\}_{2015}$ | 13.86% | 12.15% | 13.44% | 15.12% |
| $\{\text{Connections}\}_{2016}$ | 16.40% | 15.55% | 16.61% | 18.24% |
| $\{\text{Connections}\}_{2017}$ | 19.90% | 21.28% | 21.76% | 22.98% |
| $\{\text{Connections}\}_{2018}$ | 23.47% | 28.43% | 28.14% | 28.63% |

**Notes:** This table reports the aggregate output gains *relative* to the no agency counterfactual where we use the predicted labor flows with the agency network as in 2011-2018.

# B Proofs

This appendix provides the proofs for the model in Section 4. Throughout, assume $0 < \alpha < 1$, $0 < \beta < 1$, $\alpha + \beta < 1$. Let

$$Z_r(K_r) \equiv A_r K_r^{\alpha} \overline{D}_r^{1-\alpha-\beta}.$$

The production function without agency workers can be written as $Y_r = Z_r(K_r)\left(L_{Ur}^{\rho} + L_{Sr}^{\rho}\right)^{\beta/\rho}$. When agency workers are used in region $R$, effective unskilled labor is $H_{UR} = \left(L_{UR}^{\eta} + \zeta L_a^{\eta}\right)^{1/\eta}$, and output in region $R$ is $Y_R^a = Z_R(K_R)\left(H_{UR}^{\rho} + L_{SR}^{\rho}\right)^{\beta/\rho}$.

## B.1 Firm's Optimal Labor Demand with Agency Workers

Given $K_R$, $L_a$, $p_a$, and local wages $(w_R^U, w_R^S)$, the firm in region $R$ solves

$$\tilde{\pi}_R(K_R, p_a, L_a) = \max_{L_{UR}, L_{SR}} \left\{ Z_R(K_R) \left( H_{UR}^{\rho} + L_{SR}^{\rho} \right)^{\beta/\rho} - w_R^U L_{UR} - w_R^S L_{SR} - p_a L_a \right\},$$

where

$$H_{UR} = \left( L_{UR}^{\eta} + \zeta L_a^{\eta} \right)^{1/\eta}.$$

For an interior solution with respect to local unskilled and skilled labor, the first-order conditions are

$$Z_R(K_R)\beta \left( H_{UR}^{\rho} + L_{SR}^{\rho} \right)^{\frac{\beta}{\rho}-1} H_{UR}^{\rho-\eta} L_{UR}^{\eta-1} = w_R^U, \tag{34}$$

$$Z_R(K_R)\beta \left( H_{UR}^{\rho} + L_{SR}^{\rho} \right)^{\frac{\beta}{\rho}-1} L_{SR}^{\rho-1} = w_R^S. \tag{35}$$

These two conditions characterize the demand for local workers as functions of $(K_R, L_a, w_R^U, w_R^S)$.

The marginal contribution of an additional agency worker to optimized gross operating surplus is given by the envelope theorem. Define

$$\Pi_R(K_R, L_a) \equiv \max_{L_{UR}, L_{SR}} \left\{ Z_R(K_R) \left( H_{UR}^{\rho} + L_{SR}^{\rho} \right)^{\beta/\rho} - w_R^U L_{UR} - w_R^S L_{SR} \right\}.$$

Then

$$\frac{\partial \Pi_R(K_R, L_a)}{\partial L_a} = Z_R(K_R)\beta \left( H_{UR}^{\rho} + L_{SR}^{\rho} \right)^{\frac{\beta}{\rho}-1} H_{UR}^{\rho-\eta} \zeta L_a^{\eta-1}. \tag{36}$$

This is the marginal willingness to pay for agency labor before the payment $p_a$ to agencies. The firm hires agency labor only if the contract price $p_a$ is weakly below this marginal willingness to pay at the contracted quantity. In equilibrium, $p_a$ will be determined by bargaining between firms and agencies.

## B.2 Frictionless Labor Allocation without Agencies

We first characterize the labor allocation that would obtain in the absence of mobility barriers and agency activity. This is the allocation denoted by $\hat{L}_{srt}$ in the main text.

In a frictionless allocation, workers of each skill type move until skill-specific wages are equalized across regions. The first-order conditions for labor demand imply

$$w_U = Z_r(K_r)\beta \left( L_{Ur}^{\rho} + L_{Sr}^{\rho} \right)^{\frac{\beta}{\rho}-1} L_{Ur}^{\rho-1}, \tag{37}$$

$$w_S = Z_r(K_r)\beta \left( L_{Ur}^{\rho} + L_{Sr}^{\rho} \right)^{\frac{\beta}{\rho}-1} L_{Sr}^{\rho-1}. \tag{38}$$

Taking the ratio of these two equations gives

$$\frac{w_U}{w_S} = \left( \frac{L_{Ur}}{L_{Sr}} \right)^{\rho-1}.$$

Since $(w_U, w_S)$ are common across regions in the frictionless allocation, the skill ratio $L_{Ur}/L_{Sr}$ must also be common across regions. Hence each region receives the same share of the economy-wide stocks of skilled and unskilled labor:

$$\hat{L}_{Ur} = Q_r \overline{L}_U, \qquad \hat{L}_{Sr} = Q_r \overline{L}_S,$$

for some regional share $Q_r$ satisfying $\sum_r Q_r = 1$.

Substituting this structure into the labor first-order conditions, wage equalization implies

$$\frac{Q_R}{Q_P} = \left( \frac{A_R K_R^{\alpha} \overline{D}_R^{1-\alpha-\beta}}{A_P K_P^{\alpha} \overline{D}_P^{1-\alpha-\beta}} \right)^{\frac{1}{1-\beta}}.$$

Therefore,

$$\hat{L}_{srt} = \frac{\left( A_r K_{rt}^{\alpha} \overline{D}_r^{1-\alpha-\beta} \right)^{\frac{1}{1-\beta}}}{\sum_q \left( A_q K_{qt}^{\alpha} \overline{D}_q^{1-\alpha-\beta} \right)^{\frac{1}{1-\beta}}} \overline{L}_s \equiv Q_{rt} \overline{L}_s, \tag{39}$$

which is the expression used in the main text.

### B.3 Proof of Proposition 1

In steady state without agency activity, labor stocks are constant:

$$L_{srt} = L_{sr,t-1} = L_{sr}^{**}.$$

The sluggish mobility rule implies

$$L_{sr}^{**} = \nu_s^{**} \left( L_{sr}^{**} \right)^{\tau_s} \left( \hat{L}_{sr}^{**} \right)^{1-\tau_s}.$$

Rearranging,

$$\left( L_{sr}^{**} \right)^{1-\tau_s} = \nu_s^{**} \left( \hat{L}_{sr}^{**} \right)^{1-\tau_s}.$$

Therefore, for any two regions $R$ and $P$,

$$\frac{L_{sR}^{**}}{L_{sP}^{**}} = \frac{\hat{L}_{sR}^{**}}{\hat{L}_{sP}^{**}}.$$

Using (39), this gives

$$\frac{L_{sR}^{**}}{L_{sP}^{**}} = \left( \frac{A_R (K_R^{**})^{\alpha} \overline{D}_R^{1-\alpha-\beta}}{A_P (K_P^{**})^{\alpha} \overline{D}_P^{1-\alpha-\beta}} \right)^{\frac{1}{1-\beta}}, \qquad s \in \{U, S\}. \tag{40}$$

Hence skilled and unskilled labor have the same regional ratio in steady state:

$$\frac{L_{UR}^{**}}{L_{UP}^{**}} = \frac{L_{SR}^{**}}{L_{SP}^{**}} \equiv \ell.$$

Since the CES labor composite is homogeneous of degree one, it follows that

$$\frac{\left( (L_{UR}^{**})^{\rho} + (L_{SR}^{**})^{\rho} \right)^{1/\rho}}{\left( (L_{UP}^{**})^{\rho} + (L_{SP}^{**})^{\rho} \right)^{1/\rho}} = \ell.$$

In steady state, consumption is constant. The Euler equation implies that the marginal product of capital is equalized across regions:

$$\alpha A_R (K_R^{**})^{\alpha-1} \left( (L_{UR}^{**})^{\rho} + (L_{SR}^{**})^{\rho} \right)^{\beta/\rho} \overline{D}_R^{1-\alpha-\beta} = \alpha A_P (K_P^{**})^{\alpha-1} \left( (L_{UP}^{**})^{\rho} + (L_{SP}^{**})^{\rho} \right)^{\beta/\rho} \overline{D}_P^{1-\alpha-\beta}.$$

Let

$$k \equiv \frac{K_R^{**}}{K_P^{**}}.$$

Using the homogeneity of the labor composite, the capital Euler condition becomes

$$1 = \frac{A_R}{A_P} k^{\alpha-1} \ell^{\beta} \left( \frac{\overline{D}_R}{\overline{D}_P} \right)^{1-\alpha-\beta}. \tag{41}$$

From (40),

$$\ell^{1-\beta} = \frac{A_R}{A_P} k^{\alpha} \left( \frac{\overline{D}_R}{\overline{D}_P} \right)^{1-\alpha-\beta}. \tag{42}$$

Combining (41) and (42) yields

$$\ell = k.$$

Substituting $\ell = k$ into (42) gives

$$k^{1-\alpha-\beta} = \frac{A_R}{A_P} \left( \frac{\overline{D}_R}{\overline{D}_P} \right)^{1-\alpha-\beta}.$$

Therefore,

$$\frac{K_R^{**}}{K_P^{**}} = \frac{L_{UR}^{**}}{L_{UP}^{**}} = \frac{L_{SR}^{**}}{L_{SP}^{**}} = \left( \frac{A_R}{A_P} \right)^{\frac{1}{1-\alpha-\beta}} \frac{\overline{D}_R}{\overline{D}_P}.$$

This proves Proposition 1. The steady state is efficient because the mobility rule no longer distorts the long-run allocation: labor stocks coincide with the frictionless allocation, and the marginal product of capital is equalized across regions.

### B.4 Labor Allocation with Agencies but No Mobility Barriers

This subsection characterizes the hypothetical frictionless labor allocation taking agency activity $L_a$ as given. This allocation is used in the mobility rule with agencies.

Given $L_a$, region $R$ produces with effective unskilled labor

$$H_{UR} = \left( L_{UR}^{\eta} + \zeta L_a^{\eta} \right)^{1/\eta},$$

while region $P$ has $L_{UP} - L_a$ low-skilled workers available for local production. In the hypothetical frictionless allocation, skill-specific wages are equalized across regions. Therefore, the allocation

$$(\hat{L}_{UR}, \hat{L}_{UP}, \hat{L}_{SR}, \hat{L}_{SP})$$

solves the following system:

$$Z_R(K_R)\beta \left( H_{UR}^{\rho} + L_{SR}^{\rho} \right)^{\frac{\beta}{\rho}-1} H_{UR}^{\rho-\eta} L_{UR}^{\eta-1} = w_U, \tag{43}$$

$$Z_R(K_R)\beta \left( H_{UR}^{\rho} + L_{SR}^{\rho} \right)^{\frac{\beta}{\rho}-1} L_{SR}^{\rho-1} = w_S, \tag{44}$$

$$Z_P(K_P)\beta \left( (L_{UP} - L_a)^{\rho} + L_{SP}^{\rho} \right)^{\frac{\beta}{\rho}-1} (L_{UP} - L_a)^{\rho-1} = w_U, \tag{45}$$

$$Z_P(K_P)\beta \left( (L_{UP} - L_a)^{\rho} + L_{SP}^{\rho} \right)^{\frac{\beta}{\rho}-1} L_{SP}^{\rho-1} = w_S, \tag{46}$$

together with the resource constraints

$$L_{UR} + L_{UP} = \overline{L}_U, \qquad L_{SR} + L_{SP} = \overline{L}_S.$$

Equations (43)–(46) implicitly define the frictionless allocation $\hat{L}_{sr}$ taking $L_a$ as given. When $L_a = 0$, this system reduces to the frictionless allocation characterized in (39).

### B.5 Proof of Proposition 2

For simplicity we suppress time subscripts. Given $K_R$, local wages, and agency labor $L_a$, define the firm's optimized gross operating surplus before payments to agencies as

$$\Pi_R(K_R, L_a) \equiv \max_{L_{UR}, L_{SR}} \left\{ Z_R(K_R) \left( H_{UR}^{\rho} + L_{SR}^{\rho} \right)^{\beta/\rho} - w_R^U L_{UR} - w_R^S L_{SR} \right\}.$$

The firm's profit after paying agencies is

$$\tilde{\pi}_R(K_R, p_a, L_a) = \Pi_R(K_R, L_a) - p_a L_a.$$

Without agency workers, firm profit is

$$\pi_R(K_R) = \Pi_R(K_R, 0).$$

The agency's profit is

$$W(p_a, L_a) = (p_a - w_a(L_a))\, L_a - \phi(L_a) - F.$$

The Nash bargaining problem is

$$\max_{p_a, L_a} W(p_a, L_a)^{\chi} \left[ \Pi_R(K_R, L_a) - p_a L_a - \pi_R(K_R) \right]^{1-\chi}.$$

Let

$$B(p_a, L_a) \equiv \Pi_R(K_R, L_a) - p_a L_a - \pi_R(K_R)$$

denote the firm's surplus from the agency contract. For an interior solution, the first-order condition with respect to $p_a$ is

$$\chi \frac{L_a}{W(p_a, L_a)} - (1 - \chi) \frac{L_a}{B(p_a, L_a)} = 0.$$

Since $L_a > 0$, this implies

$$(1 - \chi) W(p_a^*, L_a^*) = \chi B(p_a^*, L_a^*). \tag{47}$$

This is equation (29) in the main text. It shows that the bargaining parameter $\chi$ determines the division of the static surplus between agencies and firms.

Next consider the choice of $L_a$. The total surplus generated by the agency contract is

$$W(p_a, L_a) + B(p_a, L_a) = \Pi_R(K_R, L_a) - \pi_R(K_R) - w_a(L_a) L_a - \phi(L_a) - F. \tag{48}$$

The transfer price $p_a$ cancels out. Therefore the contracted quantity $L_a$ maximizes total static surplus. The first-order condition is

$$\Pi_R'(K_R, L_a^*) = w_a(L_a^*) \left[ 1 + \frac{w_a'(L_a^*) L_a^*}{w_a(L_a^*)} \right] + \phi'(L_a^*). \tag{49}$$

This proves equation (28). The term

$$\frac{w_a'(L_a)L_a}{w_a(L_a)}$$

captures the upward-sloping supply of agency workers: recruiting additional agency workers requires increasing the agency wage for inframarginal workers as well. Hence the marginal recruitment cost exceeds the wage paid to agency workers.

Finally, (47) implies that $p_a^*$ adjusts to divide the surplus between firms and agencies. In contrast, the efficient quantity $L_a^*$ is pinned down by (49) and is independent of $\chi$, conditional on an interior solution and positive joint surplus. Therefore agency bargaining power affects the distribution of surplus, while the quantity of agency-mediated reallocation is determined by the marginal surplus condition.

### B.6 Steady State under Agencies

We now show that agencies are inactive in the steady state under the maintained interpretation that agency intermediation is valuable only because it helps arbitrage spatial wage disparities. Formally, assume that when regional low-skilled wages and marginal products are equalized, the joint surplus from agency intermediation is strictly negative for every $L_a > 0$:

$$\Pi_R(K_R, L_a) - \Pi_R(K_R, 0) - w_a(L_a)L_a - \phi(L_a) - F < 0 \quad \text{for all } L_a > 0. \tag{50}$$

Suppose, toward a contradiction, that there exists a steady state with positive agency activity, $L_a^{**} > 0$. The steady-state low-skilled mobility equations are

$$L_{UR}^{**} = \nu_U^{**} \left(L_{UR}^{**} + qL_a^{**}\right)^{\tau_U} \left(\hat{L}_{UR}^{**}\right)^{1-\tau_U}, \tag{51}$$

$$L_{UP}^{**} = \nu_U^{**} \left(L_{UP}^{**} - qL_a^{**}\right)^{\tau_U} \left(\hat{L}_{UP}^{**}\right)^{1-\tau_U}. \tag{52}$$

Since $L_a^{**} > 0$ and $q > 0$,

$$L_{UR}^{**} + qL_a^{**} > L_{UR}^{**}, \qquad L_{UP}^{**} - qL_a^{**} < L_{UP}^{**}.$$

For the two steady-state mobility equations to hold with a common normalization factor, the frictionless low-skilled allocation taking agency activity as given must satisfy

$$\hat{L}_{UR}^{**} < L_{UR}^{**}, \qquad \hat{L}_{UP}^{**} > L_{UP}^{**}.$$

In words, if agencies permanently move a positive mass of workers from $P$ to $R$ every period, then the permanent labor stock in region $R$ must lie above its frictionless benchmark, while the permanent labor stock in region $P$ must lie below its frictionless benchmark. Otherwise the mobility rule would not be stationary.

Since the marginal product of low-skilled labor is decreasing in local low-skilled labor input, this implies that the low-skilled wage in region $R$ is weakly below the frictionless equalized wage, while the low-skilled wage in region $P$ is weakly above it:

$$w_R^{U**} < \hat{w}^U < w_P^{U**}.$$

Thus a positive steady-state agency flow would require agencies to move workers from a high-wage

region $P$ to a low-wage region $R$. Under condition (50), such intermediation generates negative joint surplus for every $L_a > 0$. The Nash product cannot be positive, and agencies cannot earn nonnegative profits. This contradicts the entry condition that agencies operate only if $W(p_a^*, L_a^*) \geq 0$. Therefore, in any steady state, $L_a^{**} = 0$.

With $L_a^{**} = 0$, the economy reduces to the no-agency steady state characterized in Proposition 1. Hence agencies affect the transition path but not the efficient steady-state allocation. Their role is transitional: they operate when regional wage disparities create profitable arbitrage opportunities, and they exit once these opportunities disappear.