Agglomeration and Firm Wage Inequality: Evidence from China

Chen, Anping and Dai, Tianshi and Partridge, Mark

Jinan University, Jinan University, The Ohio State University, Jinan University, and GSSI L'Aquila, Italy

27 December 2017

Agglomeration and Firm Wage Inequality: Evidence from China

Anping Chen
Jinan University
(anping.chen@hotmail.com)

Tianshi Dai
Jinan University
(tianshidai@jnu.edu.cn)

Mark D. Partridge
The Ohio State University, Jinan University, and GSSI L'Aquila, Italy
(partridge.27@osu.edu)

December 27, 2017

Abstract

China is experiencing rapid urbanization with the steady emergence of large cities, leading to policy discussions of the role of large cities in its development. While the consensus is that agglomeration plays an important role in economic development and large cities can act as engines of economic growth, there is relatively little empirical knowledge of the effects of agglomeration on inequality. In this study, we apply panel data from a micro firm-level survey and from city-level data to investigate whether there is a causal relationship between agglomeration and establishment wage dispersion in China. Given potential endogeneity of city size, we employ an instrumental variable regression (IV) approach. We find strong evidence that agglomeration has significant effects on wage dispersion in the short- and long-run. The link between agglomeration and wage dispersion is heterogeneous across regions. The spatially varying results appear to be due to different stages of development. Our results are consistent with two-sided sorting models in that it appears that the most productive and least productive firms are moving from inland cities to the coast.

Keywords: agglomeration, wage dispersion, city size, inequality, China

Acknowledgements. We thank Qiang Li, Yi Niu and Xiwei Zhu for their helpful comments. This paper has benefited from discussions in workshops/seminars at Xiamen University, Zhejiang University, Jinan University and Hunan University, as well as at CES North America Conference and ERSA Congress. Chen acknowledges the financial support of National Natural Science Foundation of China (Grant No.71773036).

1. Introduction

China is experiencing rapid urbanization with the urban population share rising above 50% in 2011, representing the first time when more than half of its population resided in cities. Increasing numbers of large cities are emerging even as the central government has tried to control their size. There are already 15 cities which have a population larger than 10 million (OECD, 2015).[1] It is estimated that more large Chinese cities will develop during the 13th-Five Year Plan that runs from 2016 to 2020.

Cities can play an important role in economic development and can act as an engine of economic growth (Puga, 2010; Chen and Partridge, 2013; Duranton, 2015; Collier and Venables, 2017). Productivity typically rises with city size due to agglomeration effects, giving them an advantage over rural areas. Especially large cities can produce positive spillover effects by providing job opportunities for rural commuters and creating new opportunities for firms to locate in nearby small cities and rural areas (Ali et al., 2011; Duvier, 2013; Kerr and Kominers, 2015; Zheng et al., 2015; Frick and Rodriguez-Pose, 2016). However, these positive factors can be offset by congestion effects that limit commuting and other opportunities (Saito and Wu, 2016; Zhang and Kockelman, 2016).

There is a large literature that examines the effects of city size on productivity and wage premium. From this literature, some have stressed that urbanization has negative effects including the possibility that it widens income inequality. There are two ways that this may occur: through how city size affects *between* city urban wage premium or *between* city income distribution; through how city size affects *within* city wage premium or *within* city income distribution.

On the *between* city wage premium or *between* city income distribution, Combes et al. (2008) show that larger agglomeration economies are associated with the sorting of high-skilled workers to earn higher returns. Agglomeration economies also suggest that firm productivity is positively related to city size (Combes et al. 2012). Higher

[1] They include Shanghai, Guangzhou, Beijing, Shenzhen, Wuhan, Chengdu, Chongqing, Tianjin, Hangzhou, Xi'an, Changzhou, Shantou, Nanjing, Jinan, Harbin (OECD, 2015).

urban productivity and the associated increase in wages provide incentives for rural-urban migration. Tougher selection increases the returns to skills and enlarges urban income inequality (Behrens and Robert-Nicoud, 2014; Behrens et al., 2014). Baum-Snow and Pavan (2012) show that differences in returns to experience between large and small cities contribute to the observed wage inequality. Chauvin et al. (2017) find that there is strong evidence of agglomeration economies that affect earnings in developing countries including China, India, and Brazil.

On the *within* city wage premium or *within* city income distribution, most of the literature focuses on individual earnings across different city sizes. For instance, Baum-Snow and Pavan (2013) find a strong positive relationship between U.S. within-city wage inequality and city size. The notion that large cities only lead to positive sorting of the most productive firms or workers appears to be somewhat of an oversimplification. Large cities appear to attract both high- and low-skilled workers (Eeckhout et al. 2014; Accetturo et al., 2014). Florida et al. (2012) find that skills have a significant effect on city wages. D'Costa and Overman (2014) demonstrate that there is an urban premium for wage levels in Britain and work experience in cities has a positive effect over time on wage growth. Combes et al. (2015) also find that migration and agglomeration have stronger effects on wage gains for skilled natives than for the unskilled in China. Likewise, Pan et al. (2016) show that the city size–wage premium varies with difference in skills among urban residents in China.

There is also evidence that high-productivity firms and low-productivity firms coexist in cities (Forslid and Okubo, 2014, 2015). In this case, large cities may have wider wage disparities since there is more variation between the productivity of firms as well as in worker skills. Somewhat consistent with this pattern, Combes et al. (2012) show that agglomeration rather than firm selection is the main factor driving spatial productivity differences in France. Faberman and Freedman (2016) find that the returns to agglomeration diffuse within a city through firm's reallocation rather than through an increase in existing-firm productivity. However, the *within*-city firm wage distribution is much less explored in the literature. We will try to help fill this gap by examining the dispersion of firm-level wage in Chinese cities.

China is an interesting case to test the link between city size and cross-firm wage disparities. It is experiencing rapid urbanization and there are large variation in city size. There are large numbers of “big” cites as well as “small” cities to better identify causal effects. Furthermore, there are large income gaps between the coastal and inland regions with the coastal region being much wealthier. Much of these income disparities relate to regional differences in firm productivity and the associated differences in the regional distribution of firms. They date back to the origins of market reforms, and it seems reasonable to assume that coastal cities are more representative of flexible competitive labor markets than in the less-developed inland. Thus, we also can test whether the effects of city size on firm wage distribution depends on the stage of development.

Related to these firm productivity and wage differences are serious concerns regarding income inequality. The Gini coefficient was 0.465 in 2016 according to the China National Bureau of Statistics, which is similar to the corresponding estimate for the United States as reported by the U.S. Census Bureau. Some research even reports a higher Chinese Gini coefficient of 0.61 in 2011 (Xie and Zhou, 2014). Therefore, it is urgent to know whether the continuing rapid growth of large cities will further increase income inequality, which would inform current policy discussions about the optimal city-size distribution.

In this study, we begin by describing our theoretical framework based on NEG models that include “footloose-capital” with firm heterogeneity of productivity to motivate the empirical work. Most notably, the theory predicts that there exist both high- and low-productivity firms in large cities and there is a corresponding structural relationship between city size and within-city firm wage dispersion. We then apply panel data from a micro firm-level survey and from city-level data to investigate whether there is a causal relationship between Chinese city size and establishment wage dispersion. Given the possible endogeneity of city size, we use a regression-based approach to construct an IV. This kind of IV is novel in that it is akin to the industry mix variable commonly used in the literature but allows industries to have different multiplier effects.

We not only examine the short-run linkages between city size and wage dispersion but also assess their long-run relationship. We believe short-run movements mostly represent shifts in labor demand and supply. However, in the long-run, we expect agglomeration effects to dominate. We argue that some of the differences in the literature may reflect this distinction. Given that firm and household adjustments to labor market shocks may differ across regions and among sectors, we explore the effects of city size on wage dispersion between the coastal and inland regions to investigate if the level of development forges agglomeration links. Likewise, we assess this relationship for different ownership structures within these two regions and whether the development story is driven by small cities that dominate the sample. Then we consider whether the results are primarily between industry or do they apply to within industry as well. We also explore whether it is high-wage or low-wage workers who benefits from agglomeration.

A brief summary of our results is as follows. There is strong evidence that agglomeration has significant effects on wage dispersion. The relationship between city size and wage dispersion is heterogeneous across regions. In the case of coastal cities, the population coefficient is insignificant in the short run and positive and significant in the long run, consistent with agglomeration increasing wage disparities; in the inland cities, city size has significantly positive effects on wage dispersion in the short run and negative effects in the long run, being driven mainly by smaller response to labor demand shock. The differing coastal/inland region results are not due to varying private/public ownership distributions or small cities dominating the sample. Rather, these differences appear to be due to differing stages of development. The results are consistent with two-sided sorting models in that the most productive and least productive firms are moving from inland cities to the coast—reducing inland wage disparities across firms and increasing them in coastal cities in the long run.

The rest of the paper proceeds as follows. In the next section we describe our conceptual framework. Section 3 presents the empirical model and estimation implementation. Section 4 describes the data and section 5 presents the estimation results. The last section concludes.

2. Conceptual Framework

This section uses a simplified version of the Forslid and Okubo (2014, 2015) "footloose capital" NEG model to describe our conceptual framework relating city size to firm wage dispersion. These models illustrate how agglomeration affects capital movement, though they do not capture labor mobility across cities, meaning that in our context, we are illustrating equilibrium effects with no further labor mobility.[2] The key feature of these models is their incorporation of firm heterogeneity in labor productivity. It is assumed that higher productivity is associated with higher capital intensity.

Consumer Behavior

There are two cities with asymmetric population (or city size). One is a large city and another one is a small city (denoted by * below). There are two types of production factors, capital and labor. The large-city share of factors is s (assumed to be greater than 0.5) and the small city's share is 1-s. It is assumed that capital is mobile between cities but capital owners are not. Labor is mobile across sectors but not across cities. Cities differ in size, but their capital-labor ratios are identical. Each city has two sectors: agriculture and manufacturing. The agriculture sector is perfectly competitive and produces a single homogeneous good (A) using a constant returns to scale technology that only uses labor. The manufacturing sector is monopolistically competitive and produces a variety of differentiated goods (M) using an increasing returns to scale technology that employs both labor and capital.[3]

Individuals derive utility from consumption of both the agriculture good and the differentiated manufacturing good.

$$U = C_M^{\mu} C_A^{1-\mu}, \quad 0 < \mu < 1 \qquad (1)$$

where μ is the constant, C_A is consumption of the homogenous good. C_M is a consumption index of the differentiated manufacturing goods:

[2] Their models assume that wages are the same between firms but we show what happens to wage dispersion when the assumption is relaxed.

[3] The sector labels should not be interpreted literally. Agriculture refers to an industry that operates in perfect competition whereas manufacturing reflects a monopolistic competitive industry.

$$C_M = \left[\int_{k \in \Psi} c_k^{(\sigma-1)/\sigma} dk \right]^{\sigma/(\sigma-1)} \quad (2)$$

where c_k is the consumed amount of variety *k*, σ is the elasticity of substitution, and Ψ is the number of varieties available. City subscripts are suppressed for ease of notation.

Each consumer spends a share μ of his income on manufacturing goods. Utility maximization gives the demand function for a domestically produced variety *i*:

$$x_i = \frac{p_i^{-\sigma} \mu Y}{\int_{k \in \Psi} p_k^{1-\sigma} dk} \quad (3)$$

where p_k is price of variety *k* and *Y* is income in the city.

Capital ownership is assumed to be fully inter-regionally diversified, making income of each city constant and independent of the location of capital. It is assumed that total expenditure (*E*) equals total factor income. Thus,

$$E = wL + \mu E / \sigma \quad (4)$$

Without loss of generality, units are chosen so $L \equiv 1$ and wage is assumed to equal 1 (see the explanation below). The income of city *j* is equal to its share of total expenditure given by

$$Y_j = s_j E = s_j \frac{\sigma}{\sigma - \mu} \quad (5)$$

Producer Behavior

Each unit of agriculture production requires one unit of labor. The homogenous good is freely traded (zero transportation costs) and it is the numeraire by assumption:

$$P_A = w = 1 \quad (6)$$

where *w* is the uniform wage rate.

The manufactured goods use both capital and labor in production. Firms are differentiated and have a firm-specific marginal production cost α_i, which is distributed with a cumulative distribution function $F(\alpha_i)$. There is a fixed amount of national capital endowment, implying the national number of firms (*N*) is constant.

Without loss of generality, N is normalized to 1. Following Forslid and Okubo (2014), it is assumed that firms with a lower α_i have a higher capital requirement. Specifically, it is assumed that the capital requirement for a firm is given by

$$h(\alpha_i) = 2 - \alpha_i^{\gamma}, \quad \gamma > 0 \tag{7}$$

where γ is a parameter. Note that $h(\alpha_i)$ is a decreasing concave function. Firm heterogeneity in the model is simply expressed by differing $h(\alpha_i)$. The total cost (TC) for firm i is:

$$TC_i = h(\alpha_\mathrm{i})\pi_i + \alpha_i x_i \tag{8}$$

where π is the return to capital (or cost of capital), $h(\alpha_\mathrm{i})\pi_i$ is the fixed cost, and $\alpha_i x_i$ is the variable cost.

Shipping the manufactured goods between cities incurs an "iceberg" transportation cost. For one unit of good traded between cities, $\tau_{ij} > 1$ units must be shipped. The trade costs are symmetric between cities so that $\tau_{ij} = \tau_{ji} = \tau$. Profit maximization by manufacturing firms gives a constant mark-up over the marginal cost:

$$p_i = \frac{\sigma}{\sigma - 1}\alpha_i \tag{9}$$

and the export price is $p_i\tau$.

Equilibrium

The return to capital for a firm in the large city is the firm's operation profit divided by its capital stock.

$$\pi(\alpha_i) = \frac{\alpha_i^{1-\sigma}}{h(\alpha_i)(\sigma - \mu)}\mu\left(\frac{s}{\overline{\Delta}} + \frac{\phi(1-s)}{\overline{\Delta}^*}\right) \tag{10}$$

where $\phi = \tau^{1-\sigma}$, ranging between 0 and 1, represents "freeness" of trade between cities (0 is autarky and 1 is zero trade cost). The right hand side is derived from equations (3) and (5), and

$$\Delta \equiv s\int_0^1 \alpha_i^{1-\sigma} dF(\alpha) + (1-s)\phi\int_0^1 \alpha_i^{1-\sigma} dF(\alpha) \tag{11}$$

$$\bar{\Delta}^* \equiv \phi s \int_0^1 \alpha_i^{1-\sigma} dF(\alpha) + (1-s) \int_0^1 \alpha_i^{1-\sigma} dF(\alpha) \tag{12}$$

Labor stock is assumed to be sufficiently large so that the homogeneous sector pins down the wage and is active in all cities.

Forslid and Okubo (2014) show that there exists both high-productivity and low-productivity firms in large cities since the effects of α_i on the return to capital depends on the ratio of $\alpha_i^{1-\sigma}$ to $h(\alpha_i)$. Given that $h(\alpha_i)$ is concave and $\alpha_i^{1-\sigma}$ is convex, under certain conditions it will be the case that the return to capital is highest for firms with a low and a high α_i. A firm will move from a small city to a large city when:

$$\pi(\alpha_i) - \pi^*(\alpha_i) - \chi = \frac{\alpha_i^{1-\sigma}}{(2-\alpha_i^{\gamma})(\sigma-\mu)}(1-\phi)\mu\left(\frac{s}{\bar{\Delta}} - \frac{(1-s)}{\bar{\Delta}^*}\right) - \chi \geq 0 \tag{13}$$

where χ is firm moving costs. The function is U-shaped in α_i under the condition that $\sigma - 1 < \gamma$. Thus, firms at both ends of the productivity distribution will tend to move to large cities since the gains from moving are higher than the moving cost while firms in the middle of the productivity distribution will tend to locate in small cities since the gains from moving are less than moving costs.

This framework establishes a structural relationship between city size and the productivity dispersion across the city's firms. Given that a firm's wage is determined by its productivity, we expect a corresponding structural relationship between city size and within-city firm wage dispersion. Empirically, we will specify the following model:

$$WageDis = f(CitySize, X) \tag{14}$$

where the dependent variable is firm wage variation across different firms within a city, *CitySize* is the city size and *X* contains the control variables.

3. Estimation Implementation

This section describes our empirical model. We begin by showing how the short-run and long-run relationships between city size and firm wage dispersion are estimated followed by a discussion of the instruments we use to address potential endogeneity in the model. Our base model described below examines aggregate city-level wage dispersion as the dependent variable. In sensitivity analysis when we describe the empirical results, we report models that instead use *within*-industry wage dispersion (for each industry) at the city level as the dependent variable.

(1) Model specification

The dependent variable (*WageDis*) is the cross-firm wage dispersion within city measured as the standard deviation of log wage across manufacturing firms. Specifically, we specify wage dispersion in city *i* in year *t as*:

$$WageDis_{i,t} = \beta_1 CitySize_{i,t} + \gamma_1 Control_{i,t} + a_i + \delta_t + \varepsilon_{i,t} \tag{15}$$

where *CitySize* is our key explanatory variable defined as log population. a_i and δ_t are city and time (year) fixed effects respectively**.** The city fixed effects account for time-invariant omitted factors in each city that might be correlated with the explanatory variables and year fixed effects account for common national effects such as the business cycle.

The theoretical model suggests we should use firm productivity. However, we believe any productivity estimate that we could derive would be so noisy that it would be nearly useless, but we have confidence in the firm's estimate of wage and wage is often used as a proxy of productivity. Therefore, we use wage to derive the dependent variable.

We include four types of control variables. The first type includes log GDP per capita (*GDP-Per-Capita*) and the ratio of university graduates to population (*Average-Education*). We use the former to control for city productivity and the latter to control for the average education level. The second type contains industrial and ownership composition at the city level. We use the share of employment in the secondary industry (*IND2*, including manufacturing and construction) to measure industrial structure. While China's economy has been liberalizing for decades, there is still a significant share of state manufacturers (especially in the early years of the survey). Likewise, among private-sector firms, there is a key distinction between foreign-controlled firms and firms controlled by Chinese ownership. Thus, we use two variables to capture the ownership structure of manufacturing in which the first is: (1) *Public-Share*: the share of manufacturing employment of state-own enterprises (SOE) plus collective ownership enterprises (COE). We define a firm as SOE or COE if the state or collective ownership share is over 50%; (2) *Foreign-Share*: the share of manufacturing employment accounted for by foreign firms. A foreign firm is defined as a firm in which foreign shareholders have a majority ownership. The omitted share is the domestic private ownership share in each city.

We also include two sets of variable grouping to capture the characteristics of manufacturing firms. First, we use the mean of log-firm employment and the standard deviation of log employment (*lnL-mean* and *lnL-SD*). While these are primarily control variables, we expect their first-order effects to reflect that larger firms tend to pay higher wages, and thus a greater standard deviation of firm size should increase wage dispersion. The second is the mean of the log-firm capital-labor ratio and its standard deviation (*lnK/L-mean* and *lnK/L-SD*). We expect that the capital/labor ratio is positively linked to higher average human capital at the firm (increasing the average wage) and a higher capital/labor ratio implies a higher marginal product of labor, again suggesting a higher average wage. Thus, the standard deviation of the capital/labor ratio will reflect technological and human capital differences across manufactures that could cause wage dispersion.

The fourth grouping of control variables accounts for the inter-sector disparity because a wide variation in the average wage across sectors would mechanically increase overall wage dispersion in the city. Thus, we use the standard deviation of log wage between manufacturing sectors (*lnWage-ind-SD*). Specifically, manufacturing firms are grouped into 30 sectors according to their two-digit industry codes and we compute the average wage within each sector and the corresponding standard deviation of wage between sectors.

When estimating equation (15) using the standard fixed effects panel model, *annual* within-city movements of the explanatory variables are what identify the model. In terms of the city log population variable, annual within changes mainly represent some combination of labor-supply and demand shocks because annual movements around the mean are too short to identify the longer-run trend of agglomeration growth that would be less affected by random shocks. In other words, the random annual movements are dominated by the long-run cross-sectional effects, meaning that almost all of the city-size effects we are trying to capture is in the fixed effect. Thus, agglomeration economies are driven by the average population, while random variations around that mean have more transitory effects.

Because the panel model represented in equation (15) is inadequate to identify the long-run effects when the variation is mostly cross-sectional, we instead assess those long-run effects behind agglomeration by applying a two-step method. In the first step, we simply estimate the panel model shown in equation (15). Then in the second step, we use the estimated city fixed effects as the dependent variable and in a subsequent model, regress it on the average characteristics over the sample period (very similar but not exactly a between regression, because we use the fixed effect as the dependent variable). We instrument for population as in the first stage in the manner described below. This model is represented as equation (16):

$$\widehat{a}_i = \beta_2 CitySize_i + \gamma_2 Control_{\mathrm{i}} + \varepsilon_{\mathrm{i}} \tag{16}$$

in which the dependent variable is the estimated city fixed effects from equation (15). The explanatory variables are mostly the same (as described above) but averaged over the sample period. Thus, this should provide the long-run response of city wage dispersion with respect to city size as it captures long-run cross-sectional effects. Therefore, we respectively interpret β_1 in equation (15) and β_2 in equation (16) as short-run and long-run effects of city size on wage dispersion respectively.

(2) Identification and Instrumental Variable Implementation

One concern with the estimation of models above is that there could be endogeneity. On one hand, wage dispersion could affect city size, most likely via migration and sorting. Moreover, wage dispersion and city size could be affected by omitted variables, creating potential endogeneity. To assess this concern, we construct the following instrumental variable (IV).

One possible instrument is the "Bartik" (1991) instrument from shift-share analysis, which has long become the workhorse instrument to identify regional labor demand shocks:

$$IndMix_C = \sum_S EmpShare_{CS} * EmpGrow_{NS} \tag{17}$$

where $IndMix_C$ is the industry mix growth rate for city C between initial period and time T. S depicts the industry (sector). $EmpShare_{CS}$ is the employment share of sector S in city C in initial period. $EmpGrow_{NS}$ is the national employment growth rate of sector S between periods 0 and T. The instrument is simply the growth rate in the city's employment if all of its industries grew at the national growth rate. This instrument is valid if there are no offsetting labor supply responses correlated with lagged industry structure (conditional on controlling for city fixed effects and other variables that would capture supply responses).

There are two possible weaknesses of the Bartik instrument. One is that it constrains growth rates for all industries in each city to equal their respective national growth rates rather than the actual pattern that some industries have larger (smaller) effects from having larger (smaller) multiplier effects than others, which reduces the strength of the instrument in the first stage. Second, the Bartik instrument is particularly ineffective when the national trends are relatively weak compared to idiosyncratic city-based shocks, again producing a statistically weak instrument in the first stage. Indeed, we feared this could be problem for China in particular as regional growth patterns have greatly diverged, especially between the coast and inland regions. Thus, it was not too surprising that when we tried the Bartik instrument, we found that it was very weak in the first stage, with F-statistics well below the 10 rule of thumb (in fact, the F-statistics were typically below 2).

To address this weakness in the Bartik instrument, we use an alternative that is akin to the industry mix variable, but allows each industry to have differing multiplier effects. Following Detang-Dessendre et al. (2016) and their use of IV_ REGRESS, we first regress the population growth between initial period in 1990 and subsequent time T on the 1990 employment shares of 42 industries (and control variables X including provincial fixed effects) in city C:

$$PopGrow_{C1990-T} = \alpha_0 + \alpha_1 Sh_{C1} + \alpha_2 Sh_{C2} + \cdots + \alpha_{42} Sh_{C42} + \beta X_C + e_C \tag{18}$$

And then use the predicted population growth rate from equation (18) and the population in initial period to compute the population in time T:

$$Pop_{CT} = (1 + \widehat{PopGrow}_{C1990-T}) Pop_{1990} \qquad (T = 1999\text{-}2007) \tag{19}$$

in which equation (19) becomes our IV_REGRESS instrument for population. We use the initial employment share in 42 sectors in 1990, which is the earliest available detailed data on industry structure. One advantage of the IV_REGRESS instrument is that because it is based on OLS using the initial city industry shares, it should be BLUE in a statistical sense and outperform the linear Bartik instrument that is also based on initial industry shares (but with the restriction about national growth rates). Further, we use deeper lags of industry structure to avoid endogeneity as 1990 predates our sample period by several years.

4. Data

The data used in this study are from three sources: (1) China Industrial Firm Survey Data. (2) *China Statistical Yearbook for Regional Economy*. (3) China Population Census Data. The firm micro data that form the basis of the analysis are from firm-level survey over the 1999 to 2007 period, which reflect data availability.[4] The survey was conducted by China's National Bureau of Statistics. It covers all manufacturing firms which have annual revenues larger than 5 million yuan.[5] Though this eliminates some very small manufacturers, the survey represents the vast majority of them (and nearly all of national manufacturing output). We use the survey data to compute wage disparities, share of employment in public and foreign firms, and

[4] See detailed discussion on the survey in Brandt et al. (2014).

[5] The survey also covers firms in other sectors such as mining and utilities. We focus on the manufacturing sector, which is the main body of the survey.

average manufacturing characteristics in each city. We measure wage by total wage bill and employment benefits divided by the total number of employees in each firm.

We use population, GDP, number of university graduates, and employment by industry from the *China Statistical Yearbook for Regional Economy* (NBS, various issues) to compute city size, productivity, average education and industry composition. Given that the population data in the sample are *Hukou* population, which is not identical to resident population, we adjust population with an interpolation method. Specifically, we use resident population in 2000 and 2010 from the population census to compute the ratio of residents to *Hukou* in these two years and then construct the ratio for other years with interpolation and derive the resident population in these years. We apply the adjusted population in this way to measure city size. However, given that there are measurement errors by using growth of adjusted population, we use growth of *Hukou* population when constructing the IV.[6] We also use the population census data in 1990 to compute the employment share of 42 sectors in each city for the instrument construction.

Chinese cities can be divided into four types based on administrative hierarchy. They are city province, provincial capital cities, prefecture cities, and county-level cities. We include all cities except county-level cities in our sample of "cities". We drop county cities since many county-level cities have few manufacturing firms with annual revenue over 5 million yuan. Cities we study cover "total cities" as defined by NBS since many manufacturing firms are not located in the functional area (district under cities) but in the suburb. Given that boundaries in some cities have changed in the sample period, we use the definition of cities in 2007 to adjust these cities so that the definition of cities is consistent in the sample period. We have 336 prefecture-level and above regions including 283 prefecture-level cities (*Di Ji Shi*), 4 city provinces (*Zhi Xia Shi*), 17 prefectures(*Di Qu*) , 32 autonomous prefectures (*Zi*

[6] Nevertheless, we use growth of adjusted population to construct the IV and find that the results do not significantly change though the standard errors are larger as expected.

Zhi Zhou) and league (*Meng*) in our sample.[7] We use the convention of referring to them as cities to ease exposition.

Table 1 near here

Table 1 provides summary statistics for cities and wage dispersion related to our sample. There is large wage dispersion among firms and the distribution of wages varies dramatically across cities. The maximum value of the standard deviation of log wage is almost 40 times of the minimum value. The average of log population in cities is about 5.6, which converts to 3.8 million people. On average, log GDP per capita is 9.2 (which convert to 12,000 Yuan in 1999 price). Less than a quarter of employment is in the secondary industry. More than 50 percent of employment is in state-owned or collective-owned enterprises and about 10 percent is in foreign enterprises. The average of log employment in firms is about 6.5 (which converts to 1,039 employees) and the average capital/labor ratio is around 4.9. The wage inequality between sectors is substantial. The standard deviation of log wage *between* sectors is about half of that among all firms, indicating that industrial composition plays a key role in affecting city-level firm wage dispersion.

Figure 1 near here

Figure 1 plots the standard deviation of log wage against city population in all cities in the sample period. As is apparent from the graph, wage dispersion increases with city size with a slope of 0.03 at 1% significance level. This is somewhat consistent with the prediction of our theoretical framework.

5. Results

(1) Base Estimation Results

The base fixed-effect IV estimation results for the entire sample are reported in Table

[7] We re-estimate the models by only including the 283 prefecture-level cities and find that the results are robust.

2.[8] The models are reported in stages as the four groupings of control variables are added and removed from the model to illustrate robustness. At the bottom of the table, we report joint F-statistic p-values for each of the four variables groupings: city productivity & education; industry & ownership variables; firm characteristic variables; and the inter-sectoral standard deviation of wages. The Hausman endogeneity test and Wid-statistics in the first stage regression are also reported in the bottom rows in the table. The Hausman test strongly suggests that city size is endogenous, whereas that the IV_REGRESS is a strong instrument in the first-stage.

The joint F-statistic p-values indicate that the four control variable groupings are always statistically significant, except for one case of the industry-structure group in column (3). Since, the individual coefficients are not always statistically significant, this suggests that there is some multicollinearity in the four groupings.

Regarding our main variable of interest, population has a positive and statistically significant association with wage dispersion at the 1 percent level in all the estimations. The population variable is robust across a wide range of specifications. Even after controlling for city productivity, average education, industrial and ownership structure, firm characteristics and inter-sector disparities, increasing city size is associated with greater inequality in manufacturing wages. Indeed, the only case in which the population coefficient is tangibly affected is when its magnitude is reduced in column (6) when the (between) inter-sectoral wage variation variable is included—i.e., some of the positive city-size/dispersion association is related to inter-sectoral wage variation. Thus, even after controlling for inter-sectoral wage dispersion, short-run shifts in population are still positively related to wage dispersion.

Overall, this short-run finding is consistent with positive short-run agglomeration effects in that both low-wage and high-wage firms sort to larger cities, increasing wage dispersion. Yet, an alternative explanation that we prefer is that short-term shocks around the mean (controlling for year fixed effects) increase wage dispersion

[8] We rerun the estimations with OLS and find that the main results do not significantly change though the coefficients of the city size variable are smaller.

as many firms increase wages (others may decrease wages) due to labor demand shocks (which is the exogenous shock in the instrument).

Table 2 and Table 3 near here

Annual movements of population around the mean are not the best way to assess long-run effects that likely drive agglomeration externalities (especially when cross-sectional level effects of population are much more important than its transitory changes). In order to assess the long-run cross-sectional contributions of each factor on the adjustment process, we regress the estimated city fixed effects from equation (15) on the average value of each variable over the sample period—i.e., a between regression but using the estimated city fixed effects as the dependent variable. As we do for the short-run regression models, we use IV in these estimations, though the Hausman tests indicate that endogeneity is not a serious concern.

The resulting long-run effects of city size on wage dispersion for the entire sample are reported in Table 3. As before, the city-size coefficient is robust across all models, while the other control variables are generally jointly significant when considering each control variable grouping, as shown by the corresponding F-statistic p-values at the bottom of the table. Regarding our key finding, the results suggest that agglomeration is negatively related to wage dispersion in the long run, which is not consistent with two-sided (productivity) sorting of firms, though it is consistent with selectivity in which low-productive firms are driven from the market in more populated cities with greater competition. Bringing the short-run and long-run results together, our initial finding is that city wage dispersion is enlarged after a shock in the short run but is negatively linked to population in the long-run.

(2) Do Coastal-Inland Developmental Differences Produce Different Patterns?

Given the large and persistent regional disparities in China, we ask whether the relationship between city size and wage dispersion is heterogeneous across regions.[9] In particular, the coastal regions are more developed and their economies were exposed to market forces and FDI at a much earlier time than the inland. On one hand, we expect that greater coastal exposure to market forces and relatively high rates of mobility due to rural-urban migration would reduce firm wage dispersion in coastal cities as competitive forces arbitrage wage differentials. On the other hand, greater coastal exposure to market forces may have produced the two-sided firm sorting that has been observed elsewhere. In order to answer this question, we divide the country into coastal and inland regions.[10] The coastal and inland panel (short-run) models are respectively reported in Table 4 and 5.[11]

Table 4 and Table 5 near here

In the coastal panel models that reflect the short-run effects in Table 4, with exceptions in column (2) and (3), population is statistically insignificant, most notably in our base models in columns (5) and (6). Conversely, for the inland results in Table 5, population is always positive and highly statistically significantly related to short-run wage dispersion. Thus, there appears to be key differences in how between-firm wage dispersion responds to short-run shocks to population. One possible explanation is that the labor market in the coastal region is well developed and labor can move quickly in response to exogenous shocks—i.e., labor can move between coastal cities to even out shocks and net flows with inland China may also change. Conversely, the less-developed inland-regional labor market is less responsive to shocks. In contrast to the migration flows that dominate the growth of coastal regions, in the inland, with a smaller response of labor to short-run demand shocks, wages must bear more of the adjustment process, in which expanding firms

---

9 For detailed discussion on China's regional disparities, see Chen and Groenewold (2013, 2014), Herrerias and Monford (2015), and Li et al. (2017).

10 We define cities in Beijing, Shanghai, Tianjin, Hebei, Shandong, Jiangsu, Zhejiang, Fujian, Guangzhou, Hainan as the coastal region and the rest as inland following Lemoine et al. (2015).

11 We note that the Hausman tests suggest that potential endogeneity is not serious for the coastal region but we still use IV estimation for the coastal region in order to compare with the other results. Nevertheless, when we rerun the regression for the coastal region with OLS, the results are consistent with those using IV.

raise their wages, increasing wage dispersion. In addition, the foreign share of local manufacturing is negatively associated with wage dispersion in the inland, which is further suggestive that as local economies open up, market forces work to narrow wage dispersion (assuming FDI is associated with more development).

Regarding the other results, Tables 4 and 5 indicate that city productivity, the public and foreign shares, capital-labor ratio, and inter-sector disparities are generally statistically significant. It is interesting that the dispersion of the capital/labor ratio is positively linked to inequality as expected, but its mean is negative and statistically significant. This pattern is inconsistent with the evidence in Piketty (2014), which finds that increasing payments to capital is the driving force of income disparities, though we caution that our results only apply to the city level and to firm averages.

Table 6 near here

The long-run between models are reported in the first two rows of Table 6, in which the other variable results are suppressed for brevity. Recall that we expect that these longer-run effects are primarily driven by agglomeration. In the case of coastal cities, the population coefficient is positive and significant in all models indicating city size positively related to wage dispersion, consistent with agglomeration increasing wage disparities.

On the contrary, city size has significantly negative effects on wage dispersion in the inland region in all models, in which the magnitude of the effects is larger than in coastal regions. One explanation is that in inland labor markets, larger cities have less wage disparities across firms because their labor markets are more developed and flexible than in small inland cities. Another explanation is consistent with the two-sided sorting models in that the most productive and least productive firms are moving from inland cities to the coast—reducing inland wage disparities across firms and increasing them in coastal cities.

Overall, these results illustrate how population shifts can be associated with very different effects in the short- and long-run. In addition, our finding how agglomeration matters in China is somewhat consistent with Chauvin et al. (2017), who argue that agglomeration is more evident in developed economies. These findings suggest that there are other channels for agglomeration to affect city development. Another practical implication is that aggregating Chinese spatial data and estimating a model can produce misleading results.[12] Because most of the observations are in the inland, it is not surprising that the aggregate results are fairly close to the inland results given the regression is producing results near the mean of the distribution. Thus, researchers should be cautious when using spatial data in China or they may draw conclusions that mainly reflect the inland, missing the key coastal region.

(3) The Role of Private and Public Ownership

Differing characteristics of private firms and SOEs may also affect wage disparities. Private firms would seemingly be more responsive to market forces. On one hand, public firms have less wage dispersion as wages are set for other reasons than just market forces. Yet, if governments use state-owned enterprises as a type of employer of last resort, then they may have higher wage dispersion in the long-run to help facilitate this purpose. Therefore, to assess whether private and public firms have differing responses to agglomeration, as well as whether different compositions of private and public firms explain the coastal/inland results, we re-estimate the regressions for the private firms and public firms respectively in coastal and inland regions.[13] The short-run panel results for the coastal private and public firms are respectively reported in Tables 7-8 and the corresponding inland results are in Tables 9-10.

Table 7 and 8 near here

The results in Table 7 show that city size is negative but insignificant in the

[12] Liu (2014) also finds differential effects across space in China though he focuses on human capital spillovers.
[13] We do not run the regression for the foreign firms since there are few foreign firms in many cities in the inland region.

estimation for coastal private firms. The share of public and foreign firms both have positive and significant effects on wage dispersion between private firms. While the private-firm mean capital/labor ratio is negatively related to city wage dispersion, the standard deviation of the capital/labor ratio is positively related. While the latter result is not necessarily surprising, the former result suggests that there may be some selection of unproductive firms out of the market, which reduces wage dispersion. Therefore, the distribution of capital intensity is correlated with wage distribution.

Most of the average firm characteristics for public firms are statistically insignificant when considered. One interesting result is that both the private-firm and public-firm inter-sectoral standard deviation is positively linked to private firm wage dispersion. This pattern appears to reflect spillovers from public firms to private firms. However, it is only the private-firm inter-sectoral standard deviation that is statistically significant, suggesting that public-sector wages do not spillover to private-firm wages.

The results for coastal public firms, which are included in Table 8, are very similar to those for coastal private firms. Again, city size is insignificant in all models. In contrast to the results for private firms, the wage gap between sectors for private firms has significant and positive effects on wage dispersion in public firms. Therefore, it seems that there are more spillovers from private firms to public firms than that from public firms to private firms in the coastal region.

Table 9 and 10 near here

The results for the inland public and private firms are similar to the entire inland sample. City size is positive and significant in all models. For private firms, city productivity has positive and significant effects on wage dispersion while education level has negative and significant effects. Most of the private-firm characteristics are significant. For public firms, city productivity also has positive and significant effects on wage dispersion but other control variables become insignificant in general. One interesting result is that in contrast to the coastal region, there are bi-direction spillovers between public firms and private firms in the inland region.

The long-run "between" effects of city size for each firm type in both regions are reported in rows 3 to 6 of Table 6. In the long run for the coastal region, the private and public population effects are positive and significant, which are similar to the coastal region as a whole in Row 1. In both the private and public samples for the inland region, population is negative and statistically related to wage dispersion. As with the coastal region, the private and public city-size results have a similar pattern as the entire inland sample results shown in row 2. Taken as a whole, the results suggest that the differing coastal/inland region results are not due to differing private/public ownership distributions.

(4) Are There Large/Small City Differences?

Given that cities in the inland region are relatively smaller than those in the coastal region, it is possible that our findings could be driven by small cities dominating the sample. In order to test this hypothesis, we use the top 150 large cities to re-estimate the models and include an interaction term between city size and dummy for inland cities (CitySize*Inland).[14] If it is really smaller cities driving the previous results, we would expect that the interaction term would be statistically insignificant. However, if our hypothesized inland-coastal story is correct, the interaction term would be significant and the population coefficient would respectively be insignificant in the short-run panel model and positive and significant in the long-run "between" model. The short run results are reported in Table 11 and the long-run results are reported in the last row of Table 6.

Table 11 near here.

The short-run panel results indicate that the main city-size coefficient is statistically insignificant and the population-inland interaction term is positive and

[14] We have 336 cities in the sample, so 150 cities are just under 50% of the sample. We also run the estimations for top 50, 100 and 200 cities, in which the results are robust.

significant. The insignificant result of the city size variable corresponds to our previous findings for the coast and the significance of the interaction term is consistent with our results for inland regions. Therefore, our findings support our hypothesis of an inland-coastal division rather than a small-large city division.

For the long-run results shown in the last row of Table 5, the main population coefficient is typically positive and significant, which is supportive of our previous coastal results. The long-run interaction terms are negative and significant, consistent with our prior long-run model findings for the inland. Thus, our conclusion is that our results are consistent with a level of development story in the coast and inland regions and they are not altered by examining large cities in isolation.

(5) Wage Dispersion within Sector

Given that inter-sectoral wage dispersion has significant effects in all the estimations above, we ask whether the relationship between city size and wage dispersion still exists if we eliminate the effects of inter-sectoral wage dispersion directly by estimating a model with *within-sector* wage dispersion being the dependent variable. We first compute the standard deviation of log wage across firms within each sector in each city and then estimate a city-sector panel data model that includes sector and city fixed effects (i.e., for every city, there is an observation for each of the individual 30 manufacturing sectors). The results are summarized in Table 12.

Table 12 about here

It is clear that the relationship between city size and *within*-sector wage dispersion is positive and significant in the short run and negative and significant in the long run for the full sample. For the coastal region, the coefficient of city size is negative but generally insignificant and positive and significant in the long run. The results for the inland region is very similar to the whole country. Thus, we conclude that the results reported in the previous sections are generally robust to alternative specifications of wage dispersion.

(6) Who Are Winners and Who Are Losers?

The wage standard deviation we use above is a good measure to show the change in wage distribution but it cannot tell us who are the relative winners/losers in the short- and long-run. To assess this, we examine wage ratios at different parts of the distribution. Specifically for each city, we calculate the firm-wage 90th percentile to the firm-wage 10th percentile ratio, the 90/50 wage ratio, and the 50/10 wage ratio. We now use these ratios as the dependent variable. The results are shown in Table 13.

Table 13 near here

The short-run panel results are reported in columns 1, 3, and 5. The first finding for the overall China results are relatively close to the inland results, which is consistent with the standard deviation results. Thus, we focus on the coastal and inland results. The coastal results are pretty clear in that across firms, population has statistically insignificant effects across all three ratios, suggesting that there are no short-term distributional effects from population shocks on the distribution of average firm wages. However, this is not the case in the inland results. In this case, short-run labor demand shocks are positively associated with the 90/10 and 50/10 ratios at the 1% significance level, but the 90/50 city-size results are nearly insignificant. Hence, positive short-term labor demand shocks tend to increase wages relatively uniformly in the upper-half of the distribution. One implication is that local demand shocks increase average wages for higher-paying inland firms, but they have smaller or even negative effects for firms at the 10th percentile. One suggestion is that this implies that lower-paying inland firms are not benefiting from favorable demand shocks.

Turning to the long-run “between” results, in columns 2, 4, and 6, the coastal distributional effects are positive and significant, suggesting that agglomeration affects the wage distribution. However, the inland results suggest that population is negatively associated with all three ratios. This pattern suggests that long-term agglomeration effects increase average wages at the top and the middle, though the top does better than the middle. They also decrease relative wages in the bottom for firms in coastal cities. On the contrary, long-term agglomeration dampens average firm wages at the top, reduces relative wages in the middle, though the middle does better than the 90th percentile. It also increases relative wages at the 10th percentile in inland cities. This result is consistent with two-sided sorting out from the largest inland cities to the largest coastal cities. If the most productive and highest paying inland firms, as well as the least-productive and lowest-paying inland firms, move to the coast, the average wage at the 10th percentile would rise in the inland and fall in the coastal region. Wages at the 90th percentile would fall in the inland and rise in the coastal region. Overall, agglomeration has very spatially distinct distributional effects on firm wages across China.

6. Conclusion

China is experiencing rapid urbanization with the steady emergence of large cities, leading to policy discussions of the role of large cities in the country's development. There is a large related literature that documents the role of city size on productivity and wage premiums. While the consensus is that agglomeration plays an important role in economic development and large cities can act as engines of economic growth, there is relatively little empirical knowledge of the effects of agglomeration on inequality. This is an important gap in the literature: given that high-productivity firms and low-productivity firms may both sort to large cities, the continuing rapid growth of large cities may further increase Chinese income inequality, which is already at high levels. Thus, there is an urgent need to understand how China's spatial development is affecting income inequality.

This study assesses these issues. We start by using a simplified version of NEG models that allow for "footloose capital" to describe our conceptual framework relating city size to firm wage dispersion. We then apply panel data from a micro firm-level survey and from city-level data to investigate whether there is a causal relationship between agglomeration and establishment wage dispersion. We not only examine the short-run linkages between city size and wage dispersion, but also assess their long-run relationship to investigate whether the effects have temporal variation. Given potential endogeneity of city size, we employ an instrumental variable regression (IV) approach.

We find strong evidence that agglomeration has positive effects on wage dispersion in the short run and negative effects in the long run for the whole country. The relationship between city size and wage dispersion is heterogeneous across regions. The inland city results are close to the overall China results, which are driven mainly by a smaller response in labor markets to demand shocks. For the coastal cities, the city size coefficient is insignificant in the short run and positive and significant in the long run, consistent with agglomeration increasing wage disparities. The differing regional results are not due to differing private/public ownership distributions. Rather, these differences appear to be due to differing stages of development. These findings suggest that there are other channels for agglomeration to affect city development and that aggregating Chinese spatial data and estimating a model can produce misleading results.

We also find that agglomeration has very spatially distinct distributional effects on wages. For the coastal cities, the long run distribution effects are positive and significant, though the short-run results were insignificant. For the inland cities, the positive short-term labor demand shocks tend to increase wages relatively uniformly in the upper-half of the distribution, while long-term agglomeration effects dampen average firm wages at the top and middle of the distribution and increase relative wages for firms at the bottom. Overall, our results are consistent with two-sided sorting models in that the most productive and least productive firms are moving from inland cities to the coast—reducing inland wage disparities across firms and increasing them in coastal cities in the long run.

Our findings indicate that the Chinese urban development strategy could be redirected if income inequality is an important concern. The government should consider both the temporal and spatial differential effects of agglomeration when making city development planning. At present most of the large cities are concentrated in the coastal region where agglomeration increases firm wage dispersion in the long run, though the short run effects are insignificant. On the contrary, wage dispersion can be reduced as city size increases in the inland region but there are too many small cities. Therefore, it could be helpful to use policy instruments to influence urbanization, especially the development of large cities in the inland region, at least for the sake of reducing income inequality.

There are at least two directions that the current research could be extended. First, we do not provide direct evidences that both the high- and low-productivity firms move to large cities as predicted by the theoretical model. It will be promising to use micro data to examine the influence of agglomeration on reallocation decisions of heterogeneous firms. Second, we focus on manufacturing firms but it is likely that the effects of agglomeration on firms in other sectors like the service sector is different. Given that the role of the service sector is increasingly important in (urban) China, it is urgent to know whether agglomeration increases or decreases wage inequality within this sector. We leave these issues for future research.

References

Accetturo, A., A. Dalamazzo and G. de Blasio, "Skill Polarization in Local Labour Markets under Share-Altering Technical Change", *Journal of Reginal Science*, 2014, 54(2): 249-272.

Ali, K., Rose M. Olfer and Mark D. Partridge, "Urban Footprints in Rural Canada: Employment Spillovers by City Size", *Regional Studies*, 2011, 45(2): 239-260.

Bartik, T.J., *Who Benefits from State and Local Economic Development Policies*? Kalamazoo, Mich.: W. E. Upjohn Institute for Employment Research. 1991.

Baum-Snow, N. and R. Pavan, "Inequality and City Size", *Review of Economics and Statistics*, 2013, 95(5): 1535-1548.

Baum-Snow, N. and R. Pavan, "Understanding the City Size Wage Gap", *Review of Economic Studies*, 2012, 79(1): 88-127.

Behrens, K., G. Duranton and F. Robert-Nicoud, "Productive Cities: Sorting Selection and Agglomeration", *Journal of Political Economy*, 2014, 122(1): 507-553.

Behrens, K. and F. Robert-Nicoud, "Survival of the Fittest in Cities: Urbanisation and Inequality", *Economic Journal*, 2014, 124: 1371-1400.

Brandt, L., J. Van Biesebroeck and Y. Zhang, "Challenges of working with the Chinese NBS firm-level data", *China Economic Review*, 2014, 30: 339-352.

Chauvin, J. P., E. Glaeser, Y. Ma and K. Tobio, "What Is Different about Urbanization in Rich and Poor Countries? Cities in Brazil, China, India and the United States", *Journal of Urban Economics*, 2017, 98:17-49.

Chen, A. and N. Groenewold, 2013, "Does Investment Allocation Affect the Inter-Regional Output Gap in China?", *China Economic Review*, 26: 197-206.

Chen, A. and N. Groenewold, 2014, "The Regional Economic Effects of A Reduction in Carbon Emission and An Evaluation of Offsetting Policies in China", *Papers in Regional Science*, 93(2): 429-453.

Chen, A. and Mark D. Partridge, "When Are Cities Engines of Growth in China? Spread and Backwash Effects across the Urban Hierarchy", *Regional Studies*, 2013, 47(8): 1313-1331.

Collier, P. and A. J. Venables, "Who Gets the Urban Surplus?", 2017, *CEPR Discussion Papers 12001.*

Combes, P., S. Demurger and S. Li, "Migration externalities in Chinese Cities", *European Economic Review*, 2015, 76: 152-167.

Combes, P., G. Duranton and L. Gobillion, "Spatial Wage Disparities: Sorting Matters!" *Journal of Urban Economics*, 2008, 63: 723–742.

Combes, P., G. Duranton, L. Gobillon and D. Puga, "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection", *Econometrica*, 2012, 80(6): 2543-2594.

D'Costa, S. and Henry G. Overman, “The Urban Wage Growth Premium: Sorting or Learning?” *Regional Science and Urban Economics*, 2014, 48: 168-179.

Détang-Dessendre, C., Mark D. Partridge, and V. Piguet, “Local Labor Market Flexibility in a Perceived Low Migration Country: The Case of French Labor Markets”, *Regional Science and Urban Economics*, 2016, 58:89-103.

Duranton, G. “Growing Through Cities in Developing Countries”, *World Bank Research Observer*, 2015, 30(1): 39-73.

Duvier, C., “Does Urban Proximity Enhance Technical Efficiency? Evidence from Chinese Agriculture”, *Journal of Regional Science*, 2013, 53(5): 923-943.

Eeckhout, J., R. Pinheiro and K. Schmidheiny, “Spatial Sorting”, *Journal of Political Economy*, 2014, 122(3): 554-620.

Faberman, R. J. and M. Freedman, “The Urban Density Premium across Establishments”, *Journal of Urban Economics*, 2016, 93, 71-84.

Florida, R., C. Mellander, K. Stolarick and A. Ross, “Cities, Skills and Wages”, *Journal of Economic Geography*, 2012, 12(2): 355-377.

Forslid, R. and T. Okubo, “Spatial Sorting with Heterogeneous Firms and Heterogeneous Sectors”, *Regional Science and Urban Economics*, 2014, 46: 42-56.

Forslid, R. and T. Okubo, “Which Firms Are Left in the Periphery? Spatial Sorting of Heterogeneous Firms with Scale Economies in Transportation”, *Journal of Regional Science*, 2015, 55(1): 51-65.

Frick, S. A. and A. Rodriguez-Pose, “Average City Size and Economic Growth”, *Cambridge Journal of Regions, Economy and Society*, 2016, 9(2): 301-318.

Herrerías, M. J. and J. O. Monfort, 2015, “Testing Stochastic Convergence across Chinese Provinces, 1952–2008”, *Regional Studies*, 49(4): 485-501.

Kerr, W. and S. Kominers, “Agglomerative Forces and Cluster Shapes”, *Review of Economics and Statistics*, 2015, 97(4): 877-899.

Lemoine, F., S. Poncet and D. Ünal, “Spatial Rebalancing and Industrial Convergence in China”, *China Economic Review*, 2015, 34: 39-63.

Li, B., T. Li, M. Yu, and B. Chen, 2017, “Can Equalization of Public Services Narrow the Regional Disparities in China? A Spatial Econometrics Approach”, *China Economic Review*, 44: 67-78.

Liu, Z., 2014, “Human Capital Externalities in Cities: Evidence from Chinese Manufacturing Firms”, *Journal of Economic Geography*, 14: 621-649.

National Bureau of Statistics, *China Statistical Yearbook for Regional Economy*, various issues, Beijing: China Statistics Press.

OECD, *OECD Urban Policy Review: China 2015*, OECD Publishing, Paris.

Pan, L., P. Mukhopadhaya and J. Li, "City Size and Wage Disparity in Segmented Labour Market in China", *Australian Economic Papers*, 2016, 55(2): 128-48.

Piketty, T., *Capital in the Twenty-First Century*, 2014, Harvard University Press.

Puga, D., "The Magnitude and Causes of Agglomeration Economies", J*ournal of Regional Science*, 2010, 50(1): 203-219.

Saito, H. and J. Wu, "Agglomeration, Congestion, and U.S. Regional Disparities in Employment Growth", 2016, *Journal of Regional Science*, 56(1): 53-71.

Xie, Y. and X. Zhou, "Income Inequality in Today's China", *PANS*, 2014, 111(19): 6928-6933.

Zhang, W. and K. M. Kockelman, "Optimal Policies in Cities with Congestion and Agglomeration Externalities: Congestion Tolls, Labor Subsidies, and Place-Based Strategies", *Journal of Urban Economics*, 2016, 95: 64-86.

Zheng, S., W. Sun, J. Wu and M. E. Kahn, "The Birth of Edge Cities in China: Measuring the Spillover Effects of Industrial Parks", 2015, *NBER Working Paper 21378*.

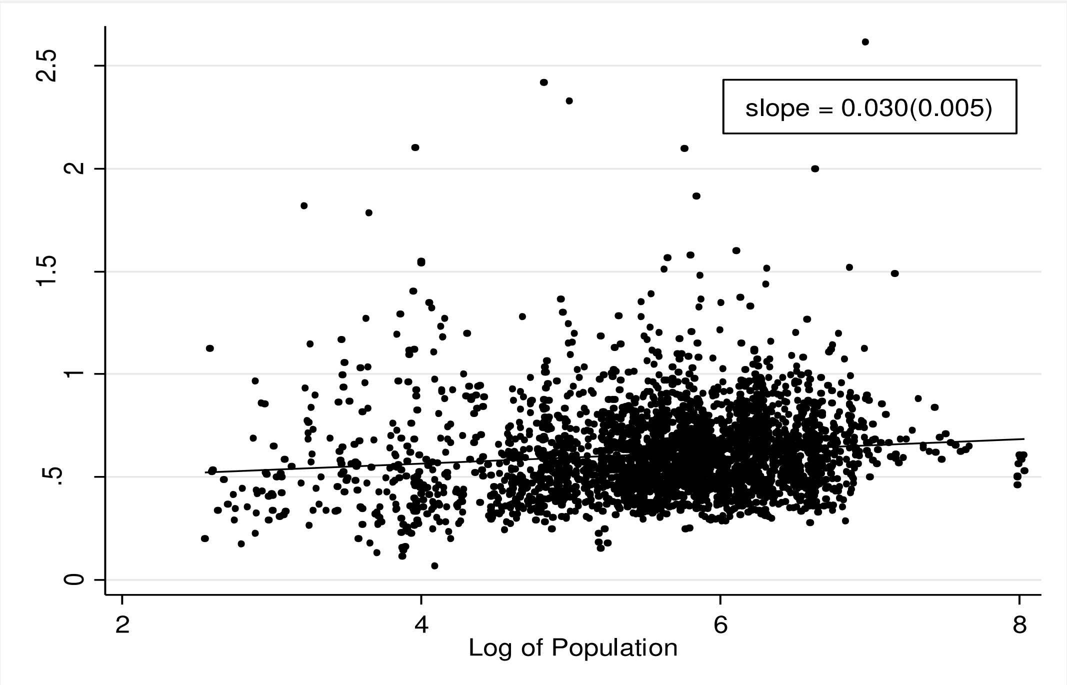

Figure 1. The correlation of within-city firm-wage dispersion and city population

Table 1: Summary Statistics of Variables

Variables	count	mean	s.d.	min	max
WageDis	3011	0.614	0.216	0.068	2.617
CitySize	3011	5.634	0.848	2.553	8.030
GDP-Per-Capita	3011	9.021	0.761	6.730	12.464
Average-Education	3011	0.001	0.003	0.000	0.031
IND2	3011	0.222	0.122	0.005	0.820
Public-Share	3011	0.535	0.280	0.000	1.000
Foreign-Share	3011	0.121	0.162	0.000	0.885
lnL-mean	3011	6.456	0.845	2.308	10.534
lnL-SD	3011	1.335	0.353	0.141	3.063
lnK/L-mean	3011	4.924	0.508	2.875	7.557
lnK/L-SD	3011	1.003	0.220	0.118	2.550
lnWage-ind-SD	3011	0.352	0.161	0.000	1.875

Notes: WageDis is the standard deviation of log wage across manufacturing firms within city. CitySize is log population. GDP-Per-Capita is log real GDP per capita. Average-Education is the ratio of university graduates to population. IND2 is the share of employment in the secondary industry. Public-Share is the share of manufacturing employment of state-own enterprises (SOE) and collective ownership enterprises (COE). Foreign-Share is the share of manufacturing employment accounted by foreign firms. lnL-mean is the mean of log firm employment. lnL-SD is the standard deviation of log firm employment. lnK/L-mean is the mean of the log firm capital-labor ratio. lnK/L-SD is the standard deviation of the log firm capital-labor ratio. LnWage-ind-SD is the standard deviation of log wage between manufacturing sectors.

Sources: China Industrial Firm Survey (NBS, various years) and *China Statistical Yearbook for Regional Economy* (NBS, various years).

Table 2: Short Run Panel Results (Full Sample)

	(1)	(2)	(3)	(4)	(5)	(6)
CitySize	0.555***	0.604***	0.592***	0.693***	0.642***	0.494***
	(0.176)	(0.187)	(0.193)	(0.214)	(0.215)	(0.170)
GDP-Per-Capita		0.149***	0.146***	0.153***	0.144***	0.121***
		(0.040)	(0.040)	(0.037)	(0.036)	(0.031)
Average-Education		-0.349	-0.316	-1.488	-1.545	-3.554*
		(2.582)	(2.565)	(2.792)	(2.535)	(1.975)
Public-Share			0.016		0.098	-0.003
			(0.065)		(0.060)	(0.039)
Foreign-Share			-0.032		-0.171	-0.240***
			(0.106)		(0.105)	(0.079)
IND2			-0.003		-0.106	-0.049
			(0.152)		(0.113)	(0.070)
lnL-mean				0.017	0.014	-0.004
				(0.025)	(0.024)	(0.014)
lnL-SD				0.021	0.017	0.001
				(0.039)	(0.040)	(0.030)
lnK/L-mean				-0.051*	-0.063**	-0.058***
				(0.029)	(0.029)	(0.021)
lnK/L-SD				0.466***	0.473***	0.244***
				(0.048)	(0.047)	(0.031)
lnWage-ind-SD						0.899***
						(0.038)
Constant	-3.395***	-5.349***	-5.235***	-6.471***	-5.859***	-4.355***
	(1.296)	(1.591)	(1.629)	(1.733)	(1.742)	(1.408)
City-fixed effects	yes	yes	yes	yes	yes	yes
Year-fixed effects	yes	yes	yes	yes	yes	yes
Observations	3011	3011	3011	3011	3011	3011
R^2	0.493	0.497	0.498	0.584	0.590	0.779
Hausman-P	0.010	0.015	0.019	0.015	0.033	0.029
Wid-Statistics	88.341	73.640	71.182	61.468	62.421	62.279
p-value-1		0.001	0.001	0.000	0.000	0.000
p-value-2			0.974		0.018	0.008
p-value-3				0.000	0.000	0.000
p-value-4						0.000

Notes: The dependent variable is the within-city standard deviation of log wage across manufacturing firms. Standard errors clustered by city are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Hausman-P is the p value of Hausman's endogeneity test statistics. Wid-Statistics is the F-statistics of IV in the first-stage regression. The last four rows report the joint significance of GDP per capita and city education, industrial and ownership structure variables, firm's characteristics variables, and inter-sector wage disparities respectively. Specifically, p-value-1 is the p value of joint significance of GDP-Per-Capita and Average-Education, p-value-2 is the p value of joint significance of Public-Share, Foreign-Share and IND2, p-value-3 is the p value of joint significance of lnL-mean, lnL-SD, lnK/L-mean and lnK/L-SD, and p-value-4 is the p value of significance of lnWage-ind-SD.

Table 3: Long Run Results (Full Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CitySize	-0.484***	-0.528***	-0.496***	-0.704***	-0.625***	-0.608***	-0.444***
	(0.036)	(0.045)	(0.043)	(0.010)	(0.009)	(0.009)	(0.006)
GDP-Per-Capita		-0.165***	-0.147***	-0.217***	-0.143***	-0.132***	-0.081***
		(0.013)	(0.024)	(0.011)	(0.017)	(0.016)	(0.011)
Average-Education		3.086	-4.152	6.398**	0.543	0.827	1.042
		(4.511)	(4.140)	(2.653)	(2.409)	(2.063)	(1.726)
Public-Share			0.291***		0.200***	0.011	0.066*
			(0.056)		(0.043)	(0.046)	(0.036)
Foreign-Share			0.131**		0.124**	-0.061	0.097**
			(0.065)		(0.051)	(0.057)	(0.048)
IND2			0.065		0.056	0.049	0.023
			(0.206)		(0.080)	(0.073)	(0.055)
lnL-mean				-0.024	-0.038***	-0.040***	-0.020***
				(0.015)	(0.014)	(0.013)	(0.008)
lnL-SD				0.062*	0.036	0.045*	0.006
				(0.035)	(0.031)	(0.027)	(0.018)
lnK/L-mean				0.061***	0.019	0.042**	0.016
				(0.016)	(0.018)	(0.019)	(0.014)
lnK/L-SD				-0.010	-0.032	-0.134***	-0.104***
				(0.042)	(0.038)	(0.039)	(0.030)
lnWage-ind-SD							-0.270***
							(0.043)
Constant	3.508***	5.512***	4.966***	6.988***	5.976***	5.912***	4.261***
	(0.212)	(0.319)	(0.232)	(0.138)	(0.148)	(0.133)	(0.109)
Provincial fixed effects	No	No	No	No	No	Yes	Yes
Observations	336	336	336	335	335	335	335
R^2	0.836	0.845	0.850	0.975	0.974	0.984	0.981
Hausman-P	0.053	0.152	0.049	0.208	0.318	0.120	0.472
Wid-Statistics	14701	26614	22607	14396	15017	9530	8049
p-value-1		0.000	0.000	0.000	0.000	0.000	0.000
p-value-2			0.000		0.000	0.650	0.112
p-value-3				0.001	0.039	0.000	0.001
p-value-4							0.000

Notes: The dependent variables are the fixed effects estimated from equation 15. The independent variables are those averaged in the sample period. Standard errors clustered by city are in parentheses . * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The first four columns correspond to the first four models in Table 2 and the fifth and sixth columns correspond to the fifth model in Table 2 and the last column correspond to the sixth column in Table 2. Hausman-P is the p value of Hausman's endogeneity test statistics. Wid-Statistics is the F-statistics of IV in the first-stage regression. The last four rows report the joint significance of GDP per capita and city education, industrial and ownership structure variables, firm's characteristics variables, and inter-sector wage disparities respectively. Specifically, p-value-1 is the p value of joint significance of GDP-Per-Capita and Average-Education, p-value-2 is the p value of joint significance of Public-Share, Foreign-Share and IND2, p-value-3 is the p value of joint significance of lnL-mean, lnL-SD, lnK/L-mean and lnK/L-SD, and p-value-4 is the p value of significance of lnWage-ind-SD.

Table 4: Short Run Panel Results (Coastal Cities)

	(1)	(2)	(3)	(4)	(5)	(6)
CitySize	-0.204	-0.344*	-0.598*	-0.065	-0.287	-0.125
	(0.147)	(0.207)	(0.309)	(0.227)	(0.279)	(0.223)
GDP-Per-Capita		0.124**	0.145**	0.109*	0.117*	0.123**
		(0.061)	(0.057)	(0.065)	(0.066)	(0.050)
Average-Education		12.519**	10.183	8.264*	6.682	2.909
		(6.178)	(6.527)	(4.700)	(4.903)	(3.342)
Public-Share			0.343**		0.352***	0.150*
			(0.136)		(0.100)	(0.083)
Foreign-Share			0.595**		0.208	0.015
			(0.240)		(0.173)	(0.142)
IND2			-0.223		-0.084	-0.010
			(0.165)		(0.137)	(0.116)
lnL-mean				0.093**	0.055	0.017
				(0.039)	(0.038)	(0.031)
lnL-SD				-0.014	0.029	-0.008
				(0.067)	(0.066)	(0.048)
lnK/L-mean				-0.097*	-0.091*	-0.090**
				(0.053)	(0.053)	(0.038)
lnK/L-SD				0.666***	0.675***	0.415***
				(0.104)	(0.106)	(0.077)
lnWage-ind-SD						0.845***
						(0.105)
Constant	2.371**	2.063	3.377	-0.687	0.694	-0.007
	(1.075)	(1.589)	(2.185)	(1.728)	(2.071)	(1.636)
City-fixed effects	yes	yes	yes	yes	yes	yes
Year-fixed effects	yes	yes	yes	yes	yes	yes
Observations	783	783	783	783	783	783
R^2	0.595	0.601	0.609	0.755	0.757	0.835
Hausman-P	0.011	0.012	0.029	0.109	0.038	0.208
Wid-Statistics	24.006	19.603	16.777	11.806	12.587	12.769
p-value-1		0.018	0.014	0.045	0.078	0.031
p-value-2			0.008		0.006	0.322
p-value-3				0.000	0.000	0.000
p-value-4						0.000

Notes: The dependent variable is the within-city standard deviation of log wage across manufacturing firms. Standard errors clustered by city are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Hausman-P is the p value of Hausman's endogeneity test statistics. Wid-Statistics is the F-statistics of IV in the first-stage regression. The last four rows report the joint significance of GDP per capita and city education, industrial and ownership structure variables, firm's characteristics variables, and inter-sector wage disparities respectively. Specifically, p-value-1 is the p value of joint significance of GDP-Per-Capita and Average-Education, p-value-2 is the p value of joint significance of Public-Share, Foreign-Share and IND2, p-value-3 is the p value of joint significance of lnL-mean, lnL-SD, lnK/L-mean and lnK/L-SD, and p-value-4 is the p value of significance of lnWage-ind-SD.

Table 5: Short Run Panel Results (Inland Cities)

	(1)	(2)	(3)	(4)	(5)	(6)
CitySize	1.025***	1.111***	1.079***	1.134***	1.104***	0.835***
	(0.264)	(0.267)	(0.272)	(0.315)	(0.311)	(0.241)
GDP-Per-Capita		0.215***	0.208***	0.225***	0.214***	0.168***
		(0.065)	(0.065)	(0.061)	(0.060)	(0.052)
Average-Education		-2.574	-1.726	-2.569	-2.271	-4.463
		(3.312)	(3.256)	(3.652)	(3.336)	(2.714)
Public-Share			-0.001		0.087	-0.008
			(0.080)		(0.071)	(0.045)
Foreign-Share			-0.149		-0.255**	-0.284***
			(0.108)		(0.120)	(0.096)
IND2			0.107		-0.034	0.004
			(0.190)		(0.145)	(0.091)
lnL-mean				0.008	0.001	-0.009
				(0.030)	(0.031)	(0.017)
lnL-SD				0.029	0.021	0.016
				(0.046)	(0.048)	(0.037)
lnK/L-mean				-0.088**	-0.098***	-0.085***
				(0.037)	(0.036)	(0.026)
lnK/L-SD				0.397***	0.404***	0.190***
				(0.050)	(0.049)	(0.032)
lnWage-ind-SD						0.894***
						(0.041)
Constant	-5.354***	-7.833***	-7.615***	-8.077***	-7.750***	-5.813***
	(1.564)	(1.884)	(1.890)	(2.081)	(2.055)	(1.630)
City-fixed effects	yes	yes	yes	yes	yes	yes
Year-fixed effects	yes	yes	yes	yes	yes	yes
Observations	2228	2228	2228	2228	2228	2228
R^2	0.429	0.434	0.438	0.514	0.521	0.752
Hausman-P	0.000	0.000	0.000	0.001	0.001	0.001
Wid-Statistics	71.048	68.027	64.239	59.457	57.382	56.924
p-value-1		0.004	0.005	0.001	0.002	0.003
p-value-2			0.502		0.030	0.026
p-value-3				0.000	0.000	0.000
p-value-4						0.000

Notes: The dependent variable is the within-city standard deviation of log wage across manufacturing firms. Standard errors clustered by city are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Hausman-P is the p value of Hausman's endogeneity test statistics. Wid-Statistics is the F-statistics of IV in the first-stage regression. The last four rows report the joint significance of GDP per capita and city education, industrial and ownership structure variables, firm's characteristics variables, and inter-sector wage disparities respectively. Specifically, p-value-1 is the p value of joint significance of GDP-Per-Capita and Average-Education, p-value-2 is the p value of joint significance of Public-Share, Foreign-Share and IND2, p-value-3 is the p value of joint significance of lnL-mean, lnL-SD, lnK/L-mean and lnK/L-SD, and p-value-4 is the p value of significance of lnWage-ind-SD.

Table 6: Long Run Effects

Panel A: Coastal and Inland Division

	Model (1)	Model (2)	Model (3)	Model(4)	Model(5)	Model(5)+	Model(6)+
Coast	0.260^{***}	0.374^{***}	0.665^{***}	0.107^{***}	0.347^{***}	0.329^{***}	0.169^{***}
	(0.024)	(0.023)	(0.020)	(0.021)	(0.018)	(0.021)	(0.017)
Inland	-0.924^{***}	-0.989^{***}	-0.959^{***}	-1.143^{***}	-1.088^{***}	-1.068^{***}	-0.786^{***}
	(0.061)	(0.078)	(0.070)	(0.011)	(0.011)	(0.010)	(0.008)

Panel B: Private and Public Firms Division in Two Regions

	Model (1)	Model(2)	Model (3)	Model (4)	Model(5)	Model (6)+	Model (7)+
Coast-Private	0.097^{***}	0.095^{***}	0.294^{***}	-0.089^{***}	0.048^{***}	0.088^{***}	0.470^{***}
	(0.019)	(0.019)	(0.021)	(0.018)	(0.018)	(0.021)	(0.025)
Coast-Public	0.293^{***}	0.454^{***}	0.830^{***}	0.431^{***}	0.710^{***}	0.605^{***}	0.379^{***}
	(0.021)	(0.021)	(0.020)	(0.023)	(0.023)	(0.027)	(0.023)
Inland-Private	-0.658^{***}	-0.722^{***}	-0.800^{***}	-1.017^{***}	-1.058^{***}	-1.150^{***}	-1.012^{***}
	(0.061)	(0.082)	(0.072)	(0.014)	(0.015)	(0.013)	(0.046)
Inland-Public	-0.758^{***}	-0.816^{***}	-0.769^{***}	-0.946^{***}	-0.991^{***}	-0.934^{***}	-0.827^{***}
	(0.051)	(0.065)	(0.058)	(0.073)	(0.042)	(0.013)	(0.040)

Panel C: Large and Small City Difference

	Model (1)	Model (2)	Model (3)	Model(4)	Model(5)	Model(5)+	Model(6)+
CitySize	0.234^{***}	0.230^{***}	0.380^{***}	0.039	0.181^{***}	0.230^{***}	0.319^{***}
	(0.033)	(0.035)	(0.034)	(0.029)	(0.026)	(0.040)	(0.037)
Interaction	-1.629^{***}	-1.648^{***}	-1.693^{***}	-1.388^{***}	-1.357^{***}	-1.386^{***}	-1.134^{***}
Term	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.047)	(0.045)

Notes: The table reports coefficients and their standard errors (clustered by city) for the city size variable. The dependent variables are fixed effects estimated from equation (15). The independent variables are those averaged in the sample period. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Panel A corresponds to the models in Table 4 and 5. Panel B corresponds to the models in Table 7-10. Panel C corresponds to the model in Table 11. “+” means that provincial dummy variable is included in the estimation.

Table 7: Short Run Panel Results (Private Firms in Coastal Cities)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CitySize	-0.053	-0.064	-0.231	0.101	-0.015	-0.048	-0.424
	(0.179)	(0.206)	(0.267)	(0.283)	(0.271)	(0.242)	(0.274)
GDP-Per-Capita		0.029	0.040	-0.014	-0.019	-0.015	-0.009
		(0.066)	(0.054)	(0.070)	(0.058)	(0.058)	(0.067)
Average-Education		1.240	-0.667	1.137	-1.190	-1.664	2.968
		(3.720)	(3.530)	(3.309)	(3.305)	(3.217)	(3.583)
Public-Share			0.259^{*}		0.288^{**}	0.358^{**}	0.355^{**}
			(0.146)		(0.143)	(0.143)	(0.150)
Foreign-Share			0.546^{***}		0.518^{**}	0.539^{***}	0.329^{*}
			(0.192)		(0.217)	(0.207)	(0.199)
IND2			-0.110		-0.105	-0.074	-0.095
			(0.173)		(0.150)	(0.147)	(0.161)
lnL-mean-private				0.010	0.007	-0.001	-0.000
				(0.031)	(0.033)	(0.034)	(0.039)
lnL-SD-private				0.025	0.066	0.082	0.069
				(0.063)	(0.067)	(0.071)	(0.072)
lnK/L-mean-private				-0.061	-0.046	-0.060^{*}	-0.071^{*}
				(0.039)	(0.036)	(0.036)	(0.042)
lnK/L-SD-private				0.449^{***}	0.453^{***}	0.453^{***}	0.553^{***}
				(0.128)	(0.124)	(0.123)	(0.143)
lnL-mean-public						0.007	-0.010
						(0.020)	(0.019)
lnL-SD-public						-0.086^{*}	-0.032
						(0.049)	(0.050)
lnK/L-mean-public						0.028	-0.010
						(0.029)	(0.031)
lnK/L-SD-public						-0.016	-0.100^{*}
						(0.045)	(0.053)
lnWage-ind-SD-private							0.297^{***}
							(0.069)
LnWage-ind-SD-public							0.036
							(0.044)
Constant	1.049	0.828	1.660	-0.202	0.304	0.519	3.369
	(1.305)	(1.795)	(2.077)	(2.097)	(2.003)	(1.860)	(2.055)
City-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	783	783	783	783	783	783	659
R^2	0.457	0.457	0.466	0.545	0.556	0.559	0.641
Hausman-P	0.091	0.117	0.245	0.618	0.803	0.708	0.024
Wid-Statistics	24.006	19.603	16.777	17.737	16.711	21.548	18.804
p-value-1		0.811	0.757	0.926	0.886	0.829	0.706
p-value-2			0.004		0.037	0.007	0.035
p-value-3				0.002	0.001	0.001	0.000
p-value-4						0.388	0.397
p-value-5							0.000

Notes: The dependent variable is the within-city standard deviation of log wage across private manufacturing firms in the coastal region. Standard errors clustered by city are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Hausman-P is the p value of Hausman's endogeneity test statistics. Wid-Statistics is the F-statistics of IV in the first-stage regression. The last five rows report the joint significance of GDP per capita and city education, industrial and ownership structure variables, firm's characteristics variables, and inter-sector wage disparities respectively. Specifically, p-value-1 is the p value of joint significance of GDP-Per-Capita and Average-Education, p-value-2 is the p value of joint significance of Public-Share, Foreign-Share and IND2, p-value-3 is the p value of joint significance of lnL-mean-private, lnL-SD-private, lnK/L-mean-private and lnK/L-SD-private, p-value-4 is the p value of joint significance of lnL-mean-public, lnL-SD-public, lnK/L-mean-public and lnK/L-SD-public, and p-value-5 is the p value of joint significance of lnWage-ind-SD-private and lnWage-ind-SD-public.

Table 8: Short Run Panel Results (Public Firms in Coastal Cities)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CitySize	-0.241	-0.421	-0.759*	-0.444	-0.741	-0.559	-0.327
	(0.242)	(0.337)	(0.449)	(0.448)	(0.532)	(0.489)	(0.365)
GDP-Per-Capita		0.120*	0.147**	0.102	0.131*	0.105	0.034
		(0.072)	(0.070)	(0.084)	(0.080)	(0.075)	(0.079)
Average-Education		15.531**	13.398*	12.385*	9.268	7.758	5.875
		(6.647)	(6.866)	(7.020)	(7.066)	(6.693)	(5.539)
Public-Share			0.404***		0.427***	0.408***	0.280**
			(0.151)		(0.130)	(0.140)	(0.127)
Foreign-Share			0.494**		0.495**	0.478**	0.307*
			(0.212)		(0.208)	(0.206)	(0.176)
IND2			-0.282		-0.218	-0.225	-0.077
			(0.224)		(0.209)	(0.207)	(0.199)
lnL-mean-public				0.021	0.015	0.016	0.044
				(0.033)	(0.035)	(0.033)	(0.031)
lnL-SD-public				0.021	-0.001	0.002	-0.024
				(0.053)	(0.059)	(0.061)	(0.065)
lnK/L-mean-public				-0.041	-0.014	-0.025	-0.030
				(0.045)	(0.044)	(0.044)	(0.046)
lnK/L-SD-public				0.295***	0.298***	0.304***	0.336***
				(0.080)	(0.082)	(0.077)	(0.085)
lnL-mean-private						-0.000	0.020
						(0.031)	(0.036)
lnL-SD-private						-0.006	0.007
						(0.062)	(0.064)
lnK/L-mean-private						0.046	0.020
						(0.040)	(0.040)
lnK/L-SD-private						0.059	0.002
						(0.060)	(0.059)
lnWage-ind-SD-private							0.202***
							(0.060)
LnWage-ind-SD-public							0.255***
							(0.045)
Constant	2.443	2.443	4.287	2.525	3.980	2.710	1.602
	(1.766)	(2.477)	(3.035)	(3.320)	(3.742)	(3.455)	(2.623)
City-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	783	783	783	783	783	783	659
R^2	0.514	0.519	0.517	0.564	0.562	0.576	0.649
Hausman-P	0.000	0.001	0.001	0.002	0.001	0.004	0.018
Wid-Statistics	24.006	19.603	16.777	23.485	21.454	21.548	18.804
p-value-1		0.013	0.023	0.112	0.147	0.235	0.539
p-value-2			0.031		0.006	0.020	0.129
p-value-3				0.004	0.003	0.002	0.001
p-value-4						0.777	0.810
p-value-5							0.000

Notes: The dependent variable is the within-city standard deviation of log wage across public manufacturing firms in the coastal region. Standard errors clustered by city are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Hausman-P is the p value of Hausman's endogeneity test statistics. Wid-Statistics is the F-statistics of IV in the first-stage regression. The last five rows report the joint significance of GDP per capita and city education, industrial and ownership structure variables, firm's characteristics variables, and inter-sector wage disparities respectively. Specifically, p-value-1 is the p value of joint significance of GDP-Per-Capita and Average-Education, p-value-2 is the p value of joint significance of Public-Share, Foreign-Share and IND2, p-value-3 is the p value of joint significance of lnL-mean-public, lnL-SD-public, lnK/L-mean-public and lnK/L-SD-public, p-value-4 is the p value of joint significance of lnL-mean-private, lnL-SD-private, lnK/L-mean-private and lnK/L-SD-private, and p-value-5 is the p value of joint significance of lnWage-ind-SD-private and lnWage-ind-SD-public.

Table 9: Short Run Panel Results (Private Firms in Inland Cities)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CitySize	0.847**	0.984***	1.046***	1.028***	1.077***	1.201***	1.166**
	(0.343)	(0.376)	(0.394)	(0.378)	(0.398)	(0.382)	(0.463)
GDP-Per-Capita		0.214***	0.233***	0.205***	0.218***	0.241***	0.300***
		(0.069)	(0.074)	(0.071)	(0.075)	(0.081)	(0.099)
Average-Education		-16.890***	-14.972***	-14.441***	-13.672***	-14.375***	-17.537***
		(5.234)	(4.904)	(5.002)	(4.686)	(4.718)	(5.328)
Public-Share			-0.179**		-0.105	-0.145*	-0.167*
			(0.089)		(0.084)	(0.082)	(0.091)
Foreign-Share			-0.137		-0.070	-0.068	-0.171
			(0.183)		(0.180)	(0.177)	(0.156)
IND2			-0.096		-0.081	-0.088	-0.150
			(0.198)		(0.175)	(0.181)	(0.197)
lnL-mean-private				-0.013	-0.018	-0.019	-0.041*
				(0.023)	(0.022)	(0.022)	(0.022)
lnL-SD-private				0.106***	0.101***	0.097**	0.102**
				(0.037)	(0.038)	(0.038)	(0.040)
lnK/L-mean-private				-0.139***	-0.138***	-0.144***	-0.149***
				(0.032)	(0.031)	(0.030)	(0.026)
lnK/L-SD-private				0.196***	0.195***	0.197***	0.156***
				(0.050)	(0.050)	(0.051)	(0.052)
lnL-mean-public						0.020	0.020
						(0.022)	(0.022)
lnL-SD-public						-0.017	-0.020
						(0.052)	(0.051)
lnK/L-mean-public						-0.011	-0.017
						(0.027)	(0.023)
lnK/L-SD-public						-0.035	-0.055*
						(0.030)	(0.032)
lnWage-ind-SD-private							0.391***
							(0.101)
LnWage-ind-SD-public							0.129***
							(0.032)
Constant	-4.332**	-6.977***	-7.315***	-6.746***	-6.992***	-7.917***	-8.139**
	(2.028)	(2.595)	(2.706)	(2.598)	(2.692)	(2.633)	(3.311)
City-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2107	2107	2107	2107	2107	2097	1821
R^2	0.311	0.317	0.316	0.383	0.381	0.375	0.506
Hausman-P	0.013	0.018	0.015	0.011	0.009	0.001	0.003
Wid-Statistics	57.832	54.040	50.834	54.363	50.165	49.826	33.518
p-value-1		0.000	0.000	0.002	0.002	0.001	0.001
p-value-2			0.246		0.604	0.318	0.246
p-value-3				0.000	0.000	0.000	0.000
p-value-4						0.678	0.444
p-value-5							0.000

Notes: The dependent variable is the within-city standard deviation of log wage across private manufacturing firms in the inland region. Standard errors clustered by city are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Hausman-P is the p value of Hausman's endogeneity test statistics. Wid-Statistics is the F-statistics of IV in the first-stage regression. The last five rows report the joint significance of GDP per capita and city education, industrial and ownership structure variables, firm's characteristics variables, and inter-sector wage disparities respectively. Specifically, p-value-1 is the p value of joint significance of of GDP-Per-Capita and Average-Education, p-value-2 is the p value of joint significance of Public-Share, Foreign-Share and IND2, p-value-3 is the p value of joint significance of lnL-mean-private, lnL-SD-private, lnK/L-mean-private and lnK/L-SD-private, p-value-4 is the p value of joint significance of lnL-mean-public, lnL-SD-public, lnK/L-mean-public and lnK/L-SD-public, and p-value-5 is the p value of joint significance of lnWage-ind-SD-private and lnWage-ind-SD-public.

Table 10: Short Run Panel Results (Public Firms in Inland Cities)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CitySize	0.857***	0.934***	0.880***	0.946***	0.862***	1.000***	0.955***
	(0.266)	(0.275)	(0.284)	(0.289)	(0.292)	(0.345)	(0.352)
GDP-Per-Capita		0.181***	0.170***	0.194***	0.172***	0.173**	0.213**
		(0.064)	(0.062)	(0.058)	(0.055)	(0.070)	(0.085)
Average-Education		1.080	1.128	2.398	2.083	-0.047	-3.531
		(3.780)	(3.684)	(3.824)	(3.601)	(3.847)	(3.578)
Public-Share			0.044		0.133*	0.166**	0.095
			(0.080)		(0.074)	(0.075)	(0.073)
Foreign-Share			-0.010		-0.053	-0.034	-0.099
			(0.117)		(0.124)	(0.142)	(0.134)
IND2			0.199		0.173	0.092	0.025
			(0.185)		(0.166)	(0.169)	(0.143)
lnL-mean-public				-0.012	-0.027	-0.035	-0.029
				(0.022)	(0.023)	(0.023)	(0.022)
lnL-SD-public				0.026	0.028	0.023	-0.003
				(0.040)	(0.040)	(0.044)	(0.044)
lnK/L-mean-public				-0.042*	-0.042*	-0.038	-0.050**
				(0.024)	(0.024)	(0.026)	(0.025)
lnK/L-SD-public				0.302***	0.305***	0.288***	0.245***
				(0.044)	(0.043)	(0.048)	(0.051)
lnL-mean-private						0.026	0.007
						(0.019)	(0.019)
lnL-SD-private						-0.034	-0.016
						(0.036)	(0.034)
lnK/L-mean-private						0.003	-0.004
						(0.020)	(0.020)
lnK/L-SD-private						0.023	0.003
						(0.026)	(0.030)
lnWage-ind-SD-private							0.079**
							(0.040)
LnWage-ind-SD-public							0.335***
							(0.039)
Constant	-4.381***	-6.528***	-6.224***	-6.718***	-6.087***	-6.982***	-6.866***
	(1.575)	(1.951)	(1.962)	(2.052)	(2.059)	(2.407)	(2.589)
City-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2218	2218	2218	2218	2218	2097	1821
R^2	0.451	0.454	0.459	0.512	0.521	0.530	0.593
Hausman-P	0.001	0.002	0.004	0.003	0.007	0.004	0.003
Wid-Statistics	69.462	64.801	60.098	66.610	61.920	49.826	33.518
p-value-1		0.010	0.012	0.001	0.002	0.027	0.041
p-value-2			0.742		0.228	0.057	0.265
p-value-3				0.000	0.000	0.000	0.000
p-value-4						0.514	0.985
p-value-5							0.000

Notes: The dependent variable is the within-city standard deviation of log wage across public manufacturing firms in the inland region. Standard errors clustered by city are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Hausman-P is the p value of Hausman's endogeneity test statistics. Wid-Statistics is the F-statistics of IV in the first-stage regression. The last five rows report the joint significance of GDP per capita and city education, industrial and ownership structure variables, firm's characteristics variables, and inter-sector wage disparities respectively. Specifically, p-value-1 is the p value of joint significance of GDP-Per-Capita and Average-Education, p-value-2 is the p value of joint significance of Public-Share, Foreign-Share and IND2, p-value-3 is the p value of joint significance of lnL-mean-public, lnL-SD-public, lnK/L-mean-public and lnK/L-SD-public, p-value-4 is the p value of joint significance of lnL-mean-private, lnL-SD-private, lnK/L-mean-private and lnK/L-SD-private, and p-value-5 is the p value of joint significance of lnWage-ind-SD-private and lnWage-ind-SD-public.

Table 11: Full Sample with Inland-Population Interaction

	(1)	(2)	(3)	(4)	(5)	(6)
CitySize	-0.172	-0.174	-0.318	-0.014	-0.156	-0.262
	(0.200)	(0.194)	(0.216)	(0.184)	(0.208)	(0.189)
CitySize*Inland	1.637***	1.652***	1.688***	1.390**	1.352**	1.174**
	(0.612)	(0.601)	(0.586)	(0.577)	(0.536)	(0.473)
GDP-Per-Capita		0.122	0.128*	0.099	0.119*	0.128**
		(0.077)	(0.075)	(0.072)	(0.067)	(0.053)
Average-Education		-1.826	-3.368	-1.915	-3.310	-3.716
		(4.101)	(4.052)	(3.602)	(3.353)	(2.741)
Public-Share			0.169*		0.249***	0.172***
			(0.096)		(0.075)	(0.059)
Foreign-Share			0.101		-0.118	-0.172
			(0.232)		(0.173)	(0.147)
IND2			-0.202		-0.162	0.005
			(0.146)		(0.121)	(0.102)
lnL-mean				0.033	0.017	-0.010
				(0.034)	(0.032)	(0.029)
lnL-SD				0.042	0.034	0.009
				(0.065)	(0.062)	(0.050)
lnK/L-mean				-0.090*	-0.121***	-0.112***
				(0.050)	(0.045)	(0.036)
lnK/L-SD				0.585***	0.586***	0.307***
				(0.067)	(0.067)	(0.055)
lnWage-ind-SD						0.765***
						(0.080)
Constant	1.942	0.672	1.670	-0.762	0.399	1.345
	(1.479)	(1.533)	(1.656)	(1.471)	(1.593)	(1.466)
City-fixed effects	yes	yes	yes	yes	yes	yes
Year-fixed effects	yes	yes	yes	yes	yes	yes
Observations	1350	1350	1350	1350	1350	1350
R^2	0.502	0.504	0.512	0.652	0.667	0.771
Hausman-P	0.017	0.020	0.016	0.009	0.005	0.001
Wid-Statistics	7.184	6.163	6.631	5.887	6.469	6.401
p-value-1		0.280	0.203	0.354	0.158	0.031
p-value-2			0.183		0.000	0.005
p-value-3				0.000	0.000	0.000
p-value-4						0.000

Notes: The dependent variable is the within-city standard deviation of log wage across manufacturing firms. Standard errors clustered by city are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Hausman-P is the p value of Hausman's endogeneity test statistics. Wid-Statistics is the F-statistics of IV in the first-stage regression. The last four rows report the joint significance of GDP per capita and city education, industrial and ownership structure variables, firm's characteristics variables, and inter-sector wage disparities respectively. Specifically, p-value-1 is the p value of joint significance of GDP-Per-Capita and Average-Education, p-value-2 is the p value of joint significance of Public-Share, Foreign-Share and IND2, p-value-3 is the p value of joint significance of lnL-mean, lnL-SD, lnK/L-mean and lnK/L-SD, and p-value-4 is the p value of significance of lnWage-ind-SD.

Table 12: Estimation for Within-Sector Wage Dispersion

	(1)	(2)	(3)	(4)	(5)	(6)
Full Sample, Short-Run	0.354***	0.391***	0.338**	0.449***	0.383***	0.351***
	(0.112)	(0.126)	(0.133)	(0.127)	(0.131)	(0.122)
Full Sample, Long-Run	-0.304***	-0.342***	-0.283***	-0.420***	-0.340***	-0.308***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.005)	(0.005)
Coastal, Short-Run	-0.179	-0.244	-0.480**	-0.110	-0.356*	-0.304*
	(0.111)	(0.157)	(0.232)	(0.130)	(0.184)	(0.169)
Coastal, Long-Run	0.239***	0.297***	0.553***	0.153***	0.396***	0.345***
	(0.013)	(0.013)	(0.012)	(0.011)	(0.011)	(0.011)
Inland, Short-Run	0.969***	1.072***	1.044***	1.061***	1.030***	0.930***
	(0.188)	(0.215)	(0.216)	(0.218)	(0.218)	(0.206)
Inland, Long-Run	-0.917***	-1.012***	-0.985***	-1.028***	-0.983***	-0.884***
	(0.010)	(0.013)	(0.012)	(0.007)	(0.006)	(0.006)
Sector-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
City-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports coefficients and their standard errors for the city size variable. The dependent variable is the within-sector standard deviation of log wage across firms in each city. Standard errors clustered by city are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Model (1) includes only city size, model (2) adds GDP per capita and education, model (3) further adds industrial and ownership composition, model (4) includes city size, GDP per capita, education and firm characteristics, model (5) includes all the control variables, and model (6) further includes inter-sector wage dispersion.

Table 13: Effects of Agglomeration on Wage Distribution

	90/10		50/10		90/50	
	SR	LR	SR	LR	SR	LR
China	1.105***	-1.076***	0.841***	-0.847***	0.264*	-0.229***
	(0.373)	(0.015)	(0.279)	(0.012)	(0.145)	(0.007)
Coastal	-0.252	0.283***	-0.186	0.200***	-0.066	0.083***
	(0.358)	(0.030)	(0.282)	(0.017)	(0.156)	(0.016)
Inland	1.615***	-1.591***	1.264***	-1.273***	0.351*	-0.318***
	(0.567)	(0.020)	(0.429)	(0.016)	(0.212)	(0.009)

Notes: The table reports coefficients and their standard errors (clustered by city) for the city size variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the firm-wage 90 percentile to the firm-wage 10 percentile ratio, the 50/10 wage ratio, and 90/50 wage ratio respectively. The independent variables include city size, GDP per capita, average education, industrial and ownership structure, firm's characteristics and inter-sector wage disparities.