RIETI Discussion Paper Series 17-E-047

Urban Wage Premium Revisited: Evidence from Japanese matched employer-employee data

KONDO Keisuke
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 17-E-047
March 2017

Urban Wage Premium Revisited: Evidence from Japanese matched employer-employee data*

KONDO Keisuke†
RIETI

Abstract

Recent studies on the empirics of agglomeration economies have assumed a two-step channel of the urban wage premium in which agglomeration increases total factor productivity (TFP) and results in higher wages. Therefore, the present study empirically examines the validity of this two-step channel via TFP and investigates other channels of the urban wage premium by using matched employer-employee data in the Japanese manufacturing sector. The findings show that the standard wage regression approach used in the literature captures not only the TFP channel of the urban wage premium but also other effects (e.g., firm size effects on wage). Furthermore, when this TFP channel is quantified, the city size elasticity of wage becomes smaller than those in the existing literature. The implication of this study is that, by exploiting the features of the Japanese employment system, it is possible to obtain interesting results concerning the urban wage premium, especially in regard to basic pay and bonuses.

JEL classification: J31, R12, R23
Keywords: Urban wage premium, City size, Agglomeration, Matched employer-employee data

*I thank Carlos Roberto Azzoni, Uwe Blien, Geoffrey J.D. Hewings, Yoshitsugu Kanemoto, Daiji Kawaguchi, Keisuke Kawata, Kozo Kiyota, Hisaki Kono, Tomoya Mori, Masayuki Morikawa, Yasusada Murata, Jiro Nakamura, Vicente Royuela, Yoichi Sugita, Makoto Yano, Kazufumi Yugami, Dao-Zhi Zeng, and participants of the RIETI DP Seminar, of the Economics-Demography Workshop at Nihon University, of the 2016 Autumn Meeting of the Japanese Economic Association, of the 6th Asian Seminar in Regional Science, of the 63rd Annual North American Meetings of the Regional Science Association International, and of the Tokyo Labor Economics Workshop for their helpful comments and suggestions. Naturally, any remaining errors are my own. This study was supported by the Grant-in-Aid for Young Scientists (B) of the Japan Society for the Promotion of Science (JSPS KAKENHI Grant Number JP 15K17066). I am grateful to the Ministry of Economy, Trade and Industry (METI), the Ministry of Health, Labour and Welfare (MHLW), and the Ministry of Internal Affairs and Communications (MIC) for providing micro datasets of Census of Manufacture (METI), Economic Census for Business Activity (MIC and METI), Basic Survey on Wage Structure (MHLW), Establishment and Enterprise Census (MIC), and Economic Census for Business Frame (MIC). I am grateful to Hiromi Shimada for her generous support in applying for the use of those microdata.

†Research Institute of Economy, Trade and Industry (RIETI): 1-3-1 Kasumigaseki, Chiyoda-ku, Tokyo, 100-8901, Japan. (E-mail: kondo-keisuke@rieti.go.jp)

1 Introduction

Recent studies in the literature on urban and regional economics have focused greater attention to the urban wage premium (see Combes and Gobillon, 2015), with many of these studies that having quantified the urban wage premium and found significant agglomeration benefits on wages.[1] In order to draw meaningful policy implications, recent studies have attempted to deepen our understanding of the channels and sources of the urban wage premium.[2] However, the standard wage regression approach, which has frequently been used in the empirics of agglomeration economies, is insufficient to identify the channels of urban wage premium. Therefore, the present study sheds light on the identification issues regarding the channels of the urban wage premium in the existing literature.

It is well-known that wages in urban areas are higher than those in rural areas. However, the simple use of positive variation between regional average wage and city size does not necessarily help to quantify the urban wage premium as one of the benefits from agglomeration economies. For example, larger cities attract more skilled workers and thus the geographical distribution of skills is not uniform within the economy, which also leads to a positive variation between regional average wage and city size. In this sense, one of the most essential aspects in an empirical analysis is to control for the geographical distribution of workers' characteristics, which requires the microdata of individual workers.

One of the seminal papers on the urban wage premium is Combes et al. (2008), who distinguish the channels of the urban wage premium between local advantages and workers' skills. A standard theoretical implication of wage determination is that a wage is determined at the point at which marginal revenue (price × marginal productivity of labor) equals to marginal cost (wage). Therefore, an essential view of higher wages in larger cities is how agglomeration economies locally increase the marginal productivity of workers in such locations. One of the channels of the urban wage premium is that agglomeration, on average, increases firm productivity at the local level and results in higher wages in larger cities.[3] Another channel is that agglomeration makes workers more productive

[1]See, for example, Glaeser and Maré (2001), Mion and Naticchioni (2009), Matano and Naticchioni (2012), Andersson et al. (2014), Matano and Naticchioni (2016), and Duranton (2016).

[2]It should also be emphasized that this study contributes to the labor economics literature. One of the main concerns in this literature is to explain wage determination, for example, in terms of the rate of return to education, gender inequality, labor institution, and policies. On the other hand, urban and regional economists have mainly attempted to clarify regional heterogeneities in wage distribution.

[3]Combes et al. (2008) offer a simple theoretical model concerning wages and agglomeration economies. In their settings, individual workers receive higher wages through firm productivity. At the same time, agglomeration economies play a crucial role in increasing productivity, as positive externalities. Consequently, the urban wage premium in their theoretical framework is observed as a spatial variation across local labor markets.

through their valuable experiences in the cities, which also increases wages. One key idea of Combes et al. (2008) is to control for workers' skills (the spatial sorting of skills) when identifying the former channel of the urban wage premium. Their research thus implicitly focuses on identifying the firm channel (wage ← TFP ← agglomeration).

As a new direction in this literature, de la Roca and Puga (2017) highlight the worker channel of the urban wage premium in which work experience in large cities makes workers more productive. Combes et al. (2008) consider that respective workers are *ex ante* skilled and unskilled, and their choices of location bring about uneven wage distribution(i.e., the spatial sorting of skills), which causes an overestimation of the urban wage premium. Conversely, de la Roca and Puga (2017) emphasize that, even if workers are *ex ante* identical, those in larger cities acquire skills faster than those in smaller cities. In other words, workers in larger cities become *ex post* skilled via valuable experiences, thus resulting in higher wages, as *dynamic* benefits from agglomeration economies.[4]

In line with the existing literature, the purpose of the present study is to identify the channels of the urban wage premium by utilizing matched employer–employee data. Although recent empirical studies mainly rely on the wage regression approach and use workers' panel data with interregional mobility to control for the spatial sorting of skills (e.g., Combes et al., 2008, 2011, 2010; Combes and Gobillon, 2015), this approach also captures other channels of the urban wage premium. This study also discusses the gap between theoretical explanations and empirical identification and provides a method to quantify the firm channel of the urban wage premium via TFP.

As mentioned earlier, a crucial assumption in the simple theory is that workers receive the urban wage premium through the TFP channel. However, this channel has not been directly examined. Hence, this study contributes to the literature by showing how matched employer–employee data can be applied to directly examine this linkage. In short, one key idea of this approach is to identify the TFP channel using firm information while canceling out the spatial sorting factors by taking the differences.

This study also offers interesting findings on the urban wage premium and firm productivity by exploiting the features of the Japanese employment system. A new aspect of this study is distinguishing the basic pay and bonuses from total earnings. While regular workers in Japan receive contractual cash earnings (mainly basic pay and overtime pay) on a monthly basis, they also receive bonuses twice a year (Kato, 2016).[5] Whereas monthly contractual cash earnings are rigid, the

[4]See also related studies on skills and the wage premium, such as Gould (2007), Glaeser and Resseger (2010), and Baum-Snow and Pavan (2012).

[5]See also Hashimoto (1990) and Ito (1991) for detailed explanations of the Japanese employment system.

bonuses can vary from year to year. By using these two types of earnings, the present study compares the urban wage premium between them.

The overall finding is that the standard wage regression approach used in the empirics of agglomeration economies captures not only the TFP channel of the urban wage premium but also other effects. In addition, the empirical results show that the key channel between wage and TFP is much weaker than theoretically expected. When the TFP channel of the urban wage premium is exactly quantified, the city size elasticity of wage becomes smaller than those in the existing literature. As investigated by Mion and Naticchioni (2009), the spatial sorting of establishments partly leads to an upward effect on the urban wage premium due to the spatially uneven distribution of establishment size. It is also found that large-sized establishments, which are generally located in larger cities, pay higher wages (e.g., Brown and Medoff, 1989).

Another interesting finding is that bonuses are more sensitive to city size than basic pay, since the establishment-level TFP shock affects amounts of the bonuses. Similar to the results of Carlsson et al. (2016), the present study finds that Japanese establishments flexibly adjust bonus amounts against TFP changes rather than basic pay. It is also suggested that the productivity benefits that the establishments receive are not distributed to the workers as an increase in basic pay. Thus, the urban wage premium of basic pay (hourly wage) is small, whereas that of bonuses is large.

Finally, it is worth mentioning that this approach can be extended to the literature on human capital externalities, which assumes that human capital intensity increases the productivity of firms and workers as positive externalities. As discussed in Heuermann et al. (2010), there is a strong correlation between the urban wage premium and human capital externalities, since skilled workers tend to be concentrated in large cities. For example, Acemoglu and Angrist (2001) estimate private and external rates of return to education using individual workers' and regional average years of schooling, respectively. Overall, the present study also fills the gap between the urban wage premium and human capital externalities.[6]

The remainder of this paper is organized as follows. Section 2 reviews the concept of the urban wage premium. Section 3 explains the theoretical framework concerning this premium. Section 4 explains the empirical methodology. Section 5 describes the matched employer–employee data and key variables. Section 6 discusses the estimation results. Section 7 conducts a robustness check of the findings. Finally, Section 8 presents the conclusions.

[6]The estimation results of human capital externalities are available in the Online Appendix.

2 Quantification Method of the Urban Wage Premium

A standard way to quantify the urban wage premium is based on a wage regression approach as follows:

$$\log(w_a) = \text{Const} + \theta \log(\text{CitySize}_a) + \varepsilon_a, \tag{1}$$

where w_a is the wage in city a, Const is a constant term, CitySize_a is some variable on city size (e.g., population, population density, employment density), and ε_a is an error term. Holding other things constant, the urban wage premium between cities a and b can be computed as

$$\frac{w_a - w_b}{w_b} = \left(\frac{\text{CitySize}_a}{\text{CitySize}_b}\right)^{\hat{\theta}} - 1, \tag{2}$$

where the term on the left hand side indicates the spatial percentage change in wage. In other words, the urban wage premium can be expressed as the function of the ratio of the city size between any two cities, with the key parameter beig the city size elasticity of wage θ.[7]

A numerical example also helps us understand how to calculate the urban wage premium. For example, when the ratio of the size of cities between a and b is two and $\hat{\theta} = 0.03$, the urban wage premium is approximately 2.1% ($\approx 2^{0.03} - 1$). This means that wages in city a are, on average, 2.1% higher than those in city b. As such, the main purpose of previous studies on the urban wage premium is to estimate the city size elasticity of wage.

3 Theoretical Explanation

3.1 TFP Channel of the Urban Wage Premium in the Literature

This section basically follows the partial equilibrium framework suggested by Combes et al. (2008), Combes et al. (2010), and Combes and Gobillon (2015). The profit of a representative establishment operating in area a at time t is given by

$$\pi_{at} = p_{at}q_{at} - w^{\ell}_{at}\ell_{at} - r_{at}k_{at},$$

where p_{at} is the market price of the product in area a, q_{at} is the output, ℓ_{at} is the amount of labor supply measured in effective labor ($\ell_{at} = \sum_{i\in(t)} s_{iat}l_{iat}$), w^{ℓ}_{at} is the wage rate to one unit of effective

[7] The city size elasticity of wage θ must be estimated from the spatial variation.

labor, and s_{iat} and l_{iat} represent the skills and labor supply of worker i. In addition, k_{at} represents the other factors of production, and r_{at} is their market price in area a.

It is assumed that the production function takes a Cobb-Douglas form (with constant returns to scale) as follows:

$$q_{at} = A_{at}\ell_{at}^{\xi}k_{at}^{1-\xi}, \quad 0 < \xi \leq 1, \tag{3}$$

where A_{at} is the TFP at the local level. Even if production technology is identical across firms, it is assumed that TFP varies across areas depending on city size. By solving profit maximization and introducing workers' heterogeneity, it is possible to obtain the following equation that describes the relationship between individual wage, TFP, and workers' skills:

$$w_{iat} = B_{at}A_{at}^{1/\xi}s_{iat}, \quad \text{where} \quad B_{at} \equiv \xi(1-\xi)^{(1-\xi)/\xi}\left(\frac{p_{at}}{r_{at}^{1-\xi}}\right)^{1/\xi}. \tag{4}$$

An important prediction from this wage equation is that, under the condition that p_{at} and r_{at} are identical across areas, the wage rate of worker i is determined in terms of TFP (A_{at}) and individual skills (s_{iat}).[8]

In the literature, an interpretation of why individual wages are higher in larger cities is based on the TFP channel through which agglomeration first increases TFP, after which TFP increases wages. As such, the empirics of agglomeration economies assume that TFP depends on local characteristics, such as population (Pop_{at}), as follows:[9]

$$\log(A_{at}) = \alpha \log(\text{Pop}_{at}) + v_{at}, \tag{5}$$

where v_{at} indicates other factors.

Taking the logarithm of Equation (4) and merging it with Equation (5), the following wage equation is derived from the partial equilibrium framework:

$$\log(w_{iat}) = \text{Const} + \frac{\alpha}{\xi}\log(\text{Pop}_{at}) + \frac{1}{\xi}v_{at} + \log(s_{iat}), \tag{6}$$

where Const is a constant term including composite parameters. Note that this wage equation takes

[8]Combes et al. (2008) also consider the case in which input and output markets are segmented between areas. In this case, spatial differences in wages not only occur through TFP, but also through the prices of the output and those of other factors.

[9]For simplicity, other local characteristics explaining TFP are not shown here, but they should be considered in an empirical analysis.

a similar form to the reduced one (1) estimated in the literature of the urban wage premium.

It is also worth discussing the connection between the urban wage premium and human capital externalities. As mentioned by Heuermann et al. (2010), the concept of human capital externalities is highly related with that of the urban wage premium because skilled workers tend to be concentrated in large cities. As introduced by Moretti (2004), instead of the city size variable in the literature on the urban wage premium, the literature on human capital externalities assumes that human capital intensity increases TFP as follows:

$$\log(A_{at}) = \lambda \text{Skill}_{at} + v_{at}, \tag{7}$$

where Skill_{at} is some measure of human capital intensity, such as the share of skilled workers or regional average years of schooling, in city a at time t. As derived earlier, the wage equation includes human capital externalities as follows:

$$\log(w_{iat}) = \text{Const} + \frac{\lambda}{\xi}\text{Skill}_{at} + \frac{1}{\xi}v_{at} + \log(s_{iat}). \tag{8}$$

3.2 Revisiting the Channels of the Urban Wage Premium

A crucial theoretical assumption in the existing literature is that workers receive a wage premium via TFP. This two-step TFP channel is reflected in the composite coefficient parameter of population or human capital intensity (α/ξ or λ/ξ) in Equations (4) and (8), respectively.

In addition, the current literature on the empirics of agglomeration economies has placed greater emphasis on estimating the economic impact of the TFP channel. Combes et al. (2011) point out two sources of bias when estimating the urban wage premium. First, the spatial sorting of skilled workers affects the area mean wage, and the urban wage premium includes an upward bias when skilled workers (especially those who earn high wages) are concentrated in large cities. This bias arising from the spatial sorting of skills is called "endogenous quality of labor." Second, higher wages in larger cities attract more workers from outside cities. An upward bias arising from this reverse causality between wages and agglomeration is called the "endogenous quantity of labor." Combes et al. (2010) emphasize that the spatial sorting of skills explains about half of the urban wage premium obtained by the standard method thus far, whereas the bias arising from the endogenous quantity of labor is not as crucial in an empirical analysis.

It is important to note that the urban wage premium is also generated from the workers' side. A standard theoretical implication of wage determination is that wages are determined at the point at which marginal revenue (price × marginal productivity of labor) equals to marginal cost (wage).[10] Another channel is that agglomeration directly affects workers' productivity, which also increase marginal productivity of labor, resulting in higher wage on average. Indeed, de la Roca and Puga (2017) offers evidence on learning by working in large cities. Workers in large cities receive higher wages through their increasing productivities. Gould (2007) also shows that work experience in cities increases wages and the white-collar workers continue to receive higher wages, even after leaving the cities, thus suggesting that large cities make workers more productive through a dynamic learning mechanism. As such, the direct channel of the urban wage premium can be expressed as $s_{iat} = f(\mathrm{Pop}_{at}, \boldsymbol{X}_{it})$, where $f(\cdot)$ is some function explaining workers' skills, and $\boldsymbol{X}_{it}$ denotes the vector of other individual characteristics (e.g., gender, age, education, working years). This expression says that the city size of the work location affects human capital accumulation. It is also possible to consider this mechanism in the literature on human capital externalities.

Recent attempts have been made in the empirics of agglomeration economies to distinguish the urban wage premium between the two-step channel via TFP and the spatial sorting of skills. The major approach has been to exploit panel data of workers with interregional mobility. Controlling for the spatial sorting of skills means shutting out the direct channel on workers' skills, which are absorbed as workers' fixed effects. Therefore, Combes et al. (2011) implicitly assume that area fixed effects of individual wages capture the urban wage premium arising from the TFP channel. However, this TFP channel has not directly been examined in the existing literature.

The main objective of the present study is to examine the validity of the TFP channel in which agglomeration increases TFP and results in higher wages. By using matched employer–employee data, it is possible to examine the direct linkage between wages and TFP. Although the Japanese dataset in this study is not designed as workers' panel data, this method can be used to quantify the impact of the TFP channel.

3.3 Identification Issues on the Channels of the Urban Wage Premium

This study discusses the identification issues regarding the channels of the urban wage premium. In order to derive a wage regression model, individual workers' skills are specified as follows:

$$s_{iat} = \exp\Big(\lambda \log(\mathrm{Pop}_{at}) + \boldsymbol{X}_{it}\boldsymbol{\beta}\Big),$$

[10] If markets are imperfect, then other factors also affect the wage rate.

where, for simplicity, we do not describe the dynamic learning process in large cities.[11] It is important to note that the spatial sorting of skills is captured by λ, if the workers' fixed effects are not controlled.

Inserting individual skills s_{iat} into Equation (6) and controlling for heterogeneities in industry γ_s and year τ_t, the regression model can be derived as follows:

$$\log(w_{ijast}) = \theta \log(\text{Pop}_{at}) + \boldsymbol{X}_{it}\boldsymbol{\beta} + \gamma_s + \pi_t + u_{ijast}, \tag{9}$$

where θ (= $\alpha/\xi + \lambda$) is a composite parameter, and u_{ijast} is an error term.

One identification issue regarding the channels of the urban wage premium is that the coefficient parameter of agglomeration variable θ consists of several parameters, which are not identified by estimating the final specification. Even if fundamental assumptions differ considerably between certain models, there are similar specifications in the wage regression models. However, in order to distinguish the channels of the urban wage premium, additional identification conditions are necessary.

Combes et al. (2008) propose their original method to control for the channel of workers' skills (λ) by using the panel data of workers with interregional mobility. Controlling for workers' fixed effects, they intend to estimate the TFP channel of α/ξ. However, their identification method captures not only the urban wage premium through the TFP channel but also other effects on large cities. For example, establishments in large cities are generally large, which offer higher wages (e.g., Mion and Naticchioni, 2009). Thus, the positive correlation between firm size and large cities makes identification of the TFP channel difficult even after controlling for workers' skills.

In this regard, the present study suggests another identification strategy for the TFP channel of the urban wage premium by using matched employer–employee data. The idea is to decompose θ into each factor such as agglomeration economies and other establishment characteristics. This approach directly captures the impact of α/ξ in parameter θ by utilizing matched employer–employee data.

4 Empirical Methodology

This section describes two empirical strategies that can be used to quantify the TFP channel of the urban wage premium. One empirical issue is the mismatch regarding the unit of observations when we compare between worker-level, establishment-level, and regional variables. In order to directly compare two variables in terms of the same unit of observation, a two- or three-step approach is

[11] The estimation results including dynamic learning process in large cities are available from the Online Appendix.

applied, as shown in Combes et al. (2008, 2010). For details with regard to matching the unit of observation between the variables, see Appendix C.

The first objective is to quantify the TFP channel of the urban wage premium. Matched employer–employee data help decompose the channels of the urban wage premium. Consider the following regional-level wage regressions as follow:[12]

$$\log(\overline{w}_{at}^{\text{Type}}) = \theta^{\text{Type}} \log(\text{Pop}_{at}) + \pi_{rt}^{\text{Type}} + \varepsilon_{at}^{\text{Type}}, \tag{10}$$

where $\log(\overline{w}_{at}^{\text{Type}})$ is the area-year mean wage rate, Type $\in$ {Total1, Total2, Total3} denotes the types of channels of the urban wage premium, π_{rt} is a regional block-year dummy (47 prefectures are divided into east and west blocks), and ε_{at} is an error term.[13] Note that wages $\overline{w}_{at}^{\text{Type}}$ are averaged at the area level (see Appendix C). Moreover, in this regression, relative hourly wage is used for the minimum wage, instead of nominal wage.

There are three types of area-mean wages in Regression (10). The first regression using $\log(\overline{w}_{at}^{\text{Total1}})$ captures the total effects of the urban wage premium between wages and agglomeration, while the logarithm of hourly wage $\log(\overline{w}_{at}^{\text{Total1}})$ is the area-year mean wage controlled for individual characteristics. This regression corresponds to the traditional framework before Combes et al. (2008) proposed the spatial sorting of skills. The coefficient of the logarithm of population captures the overall effects of agglomeration economies on wages, including the spatial sorting of skills and the dynamic learning effects in cities.

The second regression using $\log(\overline{w}_{at}^{\text{Total2}})$, which is the area-year mean wage controlled for not only individual characteristics but also establishment-level TFP, also captures the total effects of urban wage premium between wages and agglomeration. However, the two-step channel via TFP is excluded from the total effects of θ^{Total1}. In other words, the difference between θ^{Total1} and θ^{Total2} captures the TFP channel (wage $\leftarrow$ TFP $\leftarrow$ agglomeration).

The third regression using $\log(\overline{w}_{at}^{\text{Total3}})$, which is the area-year mean wage controlled for not only individual characteristics but also establishment-level TFP and establishment size (employment and financial capita), also captures the total effects of the urban wage premium between wages and

[12]An important aspect for empirical analysis is to use spatial variation in this regression. One might consider a fixed-effect model to control for area fixed effect. However, this model uses within-area temporal variation and drops information on agglomeration economies (i.e., size effects).

[13]One might consider directly estimating the regression (9). One issue arising from direct estimation is that insufficient control for area-specific factors can lead to bias in parameters $\boldsymbol{\beta}$. In other words, the direct estimation method cannot control for unobservable regional heterogeneity enough. For example, regional differences in labor markets can affect the rate of return to education.

agglomeration. However, the two-step channel via TFP and establishment size effects on wages are excluded from the total effects of θ^{Total1}. In other words, the difference between θ^{Total1} and θ^{Total3} captures the TFP channel as well as establishment size effects on wages.

One advantage of this approach is that there is no need to directly control for workers' fixed effects by using panel data. It is also possible to cancel out the spatial sorting of skills by taking the differences. In addition, the quantification of the urban wage premium can be derived from the difference between θ^{Type}. In particular, the validity of the TFP channel can be estimated as follows:

$$\text{Two-Step Channel of Urban Wage Premium} = \theta^{\text{Total1}} - \theta^{\text{Total2}}.$$

One estimation issue is that Regression (10) might suffer from endogeneity issues between wages and agglomeration. Combes et al. (2011) point out that the endogenous quantity of labor can lead to bias in estimating the parameter of population. Regions with higher wages attract workers from outside those regions, thus resulting in larger cities. The endogeneity of the population is dealt with by using the instrumental variable (IV) method of estimation. In the literature, Ciccone and Hall (1996) use a long-lagged population density, after which many studies, such as Combes et al. (2010) and de la Roca and Puga (2017), confirm that it helps control for endogeneity. The present study uses population density in 1930. Following de la Roca and Puga (2017), another instrument is the logarithm of the mean altitude (km), which also affects the long-term growth of cities.[14]

Furthermore, this present study considers two types of wages by exploiting the features of the Japanese labor market.[15] While regular workers receive contractual cash earnings (i.e., basic pay and overtime pay) on a monthly basis, they receive bonuses (on average, 3.5 months of pay) twice a year (Kato, 2016). Monthly contractual cash earnings are not that flexible, but the amounts of bonuses can vary from year to year. By using annual bonus amounts, it is possible to calculate the values averaged across areas ($\bar{b}_{at}$). Therefore, this study estimates the city size elasticity of bonuses, based on Regression (10) to examine which values are more responsible for agglomeration economies.

[14]Mean altitude data at the municipality level are taken from Zaiki et al. (2005).

[15]Hashimoto (1990), Ito (1991), and Kato (2016) provide detailed explanations for the Japanese employment system.

5 Data and Variables

5.1 Matched Employer–Employee Data in the Japanese Manufacturing Sector

This study combines individual workers' data with data from their working establishments in the Japanese manufacturing sector from 1993 to 2013.[16] Individual workers' data are taken from the Basic Survey on Wage Structure (BSWS), which is conducted each year by the Ministry of Health, Labour and Welfare. This study also focuses on workers in the manufacturing sector in order to match their dataset with establishment-level panel data. Establishment-level panel data in the manufacturing sector are taken from the Census of Manufacture (1993–2010, 2012–2013) and the Economic Census for Business Activity (2011) conducted by the Ministry of Economy, Trade and Industry (METI) each year.

The BSWS includes workers' information, such as gender, age, educational background (junior high school, high school, junior college, and university), type of employment (regular worker or non-regular worker), type of worker (production worker or non-production worker), class of position, actual number of days worked, hours worked (actual number of scheduled hours worked and actual number of overtime hours worked), earnings (monthly contractual cash earnings and annual special cash earnings), years of working for the establishment, and their working establishment information.[17] The BSWS is designed for the establishment level, which allows us to match the with other establishment datasets via firm name, location, telephone number.

The Census of Manufacture (CM) includes two forms: Form A (*Kou*), which covers establishments with 30 or more employees, and Form B (*Otsu*), which covers establishments with 29 employees or fewer. This study focuses on establishments with 30 employees or more since the data on capital stock are only available for Form A.

5.2 Variables

There are two dependent variables: hourly wage and bonuses. Hourly wage is calculated as follows. First, the total monthly earnings are calculated (contractual cash earnings in the BSWS) and second, hourly wage is calculated by dividing the total monthly earnings by the monthly actual number of hours worked. Bonuses are surveyed as annual special cash earnings received in the previous year.

---

[16] See Appendix A for details on the construction of matched data.

[17] Since educational information is limited to the regular workers in the BSWS, the present sample does not include non-regular workers. In addition, occupational career is only available for establishments with 100 employees or more. This study constructs dummies for two classes of managers.

Finally, hourly wage and bonuses are deflated by the consumer price index (2010=100), after which the spatial price difference is controlled for using the prefectural minimum wage.[18] In other words, the hourly wage and bonuses relative to the prefectural minimum wages are used in the regressions.

In line with the labor economics literature, this study controls for education level, age, years of working for the establishment, gender, class of position, and type of worker. Moreover, education levels are considered as dummies of high school graduates, junior college graduates, and university graduates (the baseline represents junior high school graduates).

As for TFP estimation, value added is used as a dependent variable, which is calculated as the total amount of production minus the total materials, fuel, and energy consumed and the subcontracting expenses for production outsourcing. In addition, labor is considered as the total annual hours worked. The dataset also includes the total annual number of workers. Using the average annual hours worked in the manufacturing sector, which is taken from the Monthly Labour Survey (Ministry of Health, Labour and Welfare), the total annual hours worked is calculated by multiplying the number of workers by the annual hours worked.[19] In addition, capital stock is measured as end-of-year book values, and all nominal values of outputs, intermediate inputs, and capital stock are deflated by each price index. The deflators of output price, input price, and investment price are constructed by the price indices available from the Bank of Japan (2011=1). Moreover, monthly price indices are yearly averaged. See Appendix B for more details regarding TFP estimation.

The regional variables are constructed from the municipal data of population censuses. In this case, the key explanatory variable is population.[20] The municipality-level populations are taken from the 1990, 1995, 2000, 2005, 2010, and 2015 Population Censuses, and the linear interpolation is implemented every five years using the percentage change in the population. One estimation issue is that the geographical units of the regional data correspond to the administrative units, not the economic areas. Workers and consumers cross municipal borders when commuting and shopping, thus creating a geographical mismatch between their residential locations and their locations of work and consumption. As a result, the simple use of population as the administrative unit suffers from border discontinuity and does not include potential people who can access such locations from

[18]Minimum wage was stipulated as the daily wage before 2002. Using the percentage change in the daily minimum wage at the prefecture level between years, it is possible to recursively estimate the hourly minimum wage until 1993.

[19]The CM began to distinguish workers into regular and non-regular workers from 2001 on. This study calculates the hours worked adjusted for regular and non-regular workers from 2001 to 2013. However, it does not separate the annual hours worked between regular and non-regular workers as two explanatory variables. The logarithm of the sum of hours worked for regular and non-regular workers is used as an explanatory variable.

[20]In this study, the estimation results of manufacturing establishments are quite similar to those of population. The estimation results are available in the Online Appendix.

outside areas.

To take into account potential people, including those from surrounding municipalities, this study calculates the total population within a d km radius from the centroid of the municipality as city size.[21] The local sum of the population for municipality a is expressed as $\sum_{b=1}^{R} I(d_{ab} < d) \cdot \text{Pop}_b$, in which R represents the number of municipalities, and $I(d_{ab} < d)$ is the indicator function that takes the value of 1 if the distance between municipalities a and b is less than d km, or 0 otherwise.[22] It also sets $d = 30$ km by considering local labor markets and commuting distance. Figure 1 illustrates the spatial unit of city size and includes the case of Chiyoda-ku, Tokyo, the population of which covers the surrounding 68 municipalities located within a 30 km radius from the centroid.

[Figure 1]

Table 1 presents the descriptive statistics of the variables, which are divided into the characteristics of the individual workers, establishments, and regions, respectively. The sample covers 3,060,323 individual workers in pooled data between 1993 and 2013, 32,169 establishments (100,388 in pooled data), and 1,512 municipalities (21,482 in pooled data). As of April 5, 2014, the number of municipalities in Japan is 1,741, and the dataset covers approximately 87% of the municipalities.

[Table 1]

6 Estimation Results

Figure 2 illustrates the correlations between the area mean wage controlled for individual characteristics and the city size. Consistent with previous findings, wage is positively correlated with city size. Figure 3 illustrates the correlations between area mean bonus controlled for individual characteristics and the city size. There is also a positive relationship between bonuses and city size.

[Figures 2–3]

Table 2 presents the estimation results regarding the city size elasticity of hourly wage. Columns (1) to (3) show the benchmark estimation results obtained by OLS. In Column (1), the city size elasticity

[21]This approach regarding the spatial unit is similar to that of de la Roca and Puga (2017). Maré and Graham (2013) also suggest a market potential type of employment density in order to to measure agglomeration.

[22]The latitudes and longitudes of municipalities are obtained using GIS software and the bilateral distances between any two municipalities are calculated by the formula suggested by Vincenty (1975).

of wage is positive and significant, with a value of 0.026.[23] This elasticity captures the total effects of city size, including the TFP channel, the establishment characteristic effects, the spatial sorting of skills, the dynamic learning effects. In Column (2), the impact of city size on wage (excluding the TFP channel) is estimated, with a value of 0.022. In other words, the difference between Columns (1) and (2) captures the urban wage premium arising from the TFP channel, with an implied elasticity of 0.004. This value is smaller than expected in the literature. For example, Combes et al. (2010) obtain the elasticity of 0.027 using the French workers' panel data. Furthermore, de la Roca and Puga (2017) obtain the elasticity of 0.025 using the Spanish workers' panel data, and D'Costa and Overman (2014) obtain the elasticity of 0.022 using the panel data of British workers. It is important to note that these three studies control for the spatial sorting of workers.

Column (3) provides an important clue to understanding why the estimates of this study are smaller than those in the existing literature. After controlling for establishment size in terms of employment and financial capital, a city size elasticity of 0.016 is obtained, which means that establishment size differences between cities can affect the estimation results of the urban wage premium. The difference between Columns (1) and (3) captures both the urban wage premium arising from the TFP channel and the establishment size effects on wages, with an implied elasticity is 0.010. The estimation results also suggest that, unlike the findings in the existing literature, the TFP channel of the urban wage premium only has a small magnitude.

Columns (4) to (7) in Table 2 offer a robustness check of the benchmark results using the IV estimation method. This study also controls for a reverse causality between wage and city size. As shown in Column (4), the IV for the logarithm of population includes the logarithm of the population density in 1930 and mean altitude. As discussed in Combes et al. (2011), the endogenous quantity of labor barely leads to a bias in Columns (5) to (7). The IV estimation results are basically the same as those of the OLS method.

Table 3 presents the estimation results of the city size elasticity of bonuses. Columns (1) to (3) show the benchmark estimation results obtained by the OLS method. In Column (1), the city size elasticity of bonuses is positive and significant, with a value of 0.059. This value is much greater than that of hourly wage. Again, the difference between Columns (1) and (2) captures the urban wage premium arising from the TFP channel, with an implied elasticity of 0.016 in the case of bonuses. This value is quite similar to those obtained in the existing literature (e.g., Combes et al., 2010; de la

[23]Morikawa (2011) shows that the density elasticity of real wage (hourly wage controlled for regional price index from the National Survey Prices) is 0.026 in Japan in 2009. However, note that population density measured in an administrative unit is used and his elasticity is not directly comparable with ours.

Roca and Puga, 2017; D'Costa and Overman, 2014). In addition, the difference between Columns (1) and (3) captures both the urban wage premium arising from the TFP channel and the establishment size effects on wages, with an implied elasticity of 0.032. The comparison between hourly wage and bonuses suggests that the urban wage premium is mainly observed for bonuses in Japan.

Columns (4) to (7) in Table 3 offer a robustness check of the benchmark results by using the IV estimation method. This study also controls for a reverse causality between bonus and city size. Unlike the relationship between hourly wage and city size, this case shows upward biases by the OLS estimates, whereas the IV estimates in Columns (5) to (7) show lower elasticities than those of the OLS method.

In sum, the findings raise an important question about the interpretation of the TFP channel of the urban wage premium. The existing literature interprets that the city size elasticity of wage captures the TFP channel of the urban wage premium by controlling for the spatial sorting of skills. However, the estimation results suggest that other firm characteristics related to city size provide an upward bias on the urban wage premium. Indeed, it is found that the TFP channel of urban wage premium is quite limited, even if agglomeration economies increase establishment TFP. This suggests that the previous studies not only capture the TFP channel but also the effects arising from establishment size differences between cities.

[Tables 2–3]

7 Robustness Checks

This section explores the reasons for the weak connection in the TFP channel, of which there are two possibilities: 1) the first channel between TFP and city size; and 2) the second channel between wage and TFP. The following measures the magnitudes of each channel.

7.1 Testing for City Size Effects on TFP

This section examines whether agglomeration, on average, increases TFP. The TFP regression is given by

$$\log\left(\overline{\text{TFP}}_{at}\right) = \alpha \log(\text{Pop}_{at}) + \pi_{rt} + \varepsilon_{at},$$

where $\overline{\text{TFP}}_{at}$ is the area-year mean TFP estimated from the establishment microdata (See Appendix C).

Table 4 presents the estimation results regarding TFP and agglomeration. The benchmark estimation results by the OLS method in Column (1) show that the city size elasticity of TFP is 0.090, which is higher than 0.04 obtained in Combes et al. (2010). Similar to the wage regression, Column (2) shows the IV estimation results, but the magnitude does not differ from that of the OLS. The estimation results in Table 4 provide evidence that the agglomeration increases TFP, thus supporting the first channel of the TFP channel.[24]

[Table 4]

7.2 Testing for the Channel between Wage and TFP

This section examines the validity of the second channel between wage and TFP in Equation (4). The theoretical model implies that the TFP elasticity of wage is $1/\xi$ in Equation (4) when TFP is treated as an exogenous variable. Taking the logarithm on both sides in Equation (4), the regression model is as follows:

$$\log(\overline{w}_{jat}) = \rho \log(\hat{A}_{jat}) + \eta_a + \psi_j + \gamma_s + \pi_t + e_{jat} \tag{11}$$

where $\log(\overline{w}_{jat})$ is the establishment-year mean hourly wage controlled for workers' skills s_{iat}, and e_{jat} is the error term. As earlier, this study examine the case of the establishment mean bonuses $\log(\overline{b}_{jat})$. Note that hourly wage and bonuses relative to the prefectural minimum wages are used. The parameter of interest is ρ, which is expected to be estimated as $1/\xi$. If ξ ranges from 0.4 to 0.9, ρ is expected to range from 1.1 to 2.5. The empirics of agglomeration economies inserts TFP determinants in Equation (5) into the Equation (4), since the dataset is generally limited to workers' data. However, the matched employer–employee data enable a direct estimation of Equation (4).

In addition, this study estimates ρ in Regression (11) from four types of variations. The benchmark estimation is simply regressing (11) without the area and the establishment fixed effects. Then, it examines how the magnitude ρ changes when adding the area and the establishment fixed effects into the regression in succession. The fourth case is the first difference estimation of regression (11).

Figure 4 shows the correlation between the establishment mean wage, establishment mean bonuses, and the establishment-level TFP. Panel (a) in Figure 4 illustrates the level-variation between hourly wage and TFP at the establishment level, between which there is a positive correlation. In turn, Panel (b) in Figure 4 illustrates the variation of their first differences. If we control for geo-

[24] Kondo (2016) also finds that agglomeration economies, rather than stronger selection in larger cities, better explain the spatial productivity differences in the Japanese manufacturing sector by using the quantile approach suggested by Combes et al. (2012).

graphic and establishment fixed factors, the observed positive correlation disappears, thus implying that unobservable omitted factors simultaneously explain wage and productivity. In turn, Panels (c) and (d) in Figure 4 illustrate the case of the establishment mean bonuses. The regression analysis examines the relationship observed in Figure 4.

[Figure 4]

Table 5 presents the estimation results of the regression analysis. Column (1) includes the benchmark results, with a TFP elasticity of hourly wage of 0.081, while Column (5) offers the benchmark results, with a TFP elasticity of bonuses of 0.237. Firm-level TFP has minimal effects on hourly wage, whereas it has greater impacts on bonuses. Although this parameter value (theoretically expected from the partial equilibrium) is more than 1, the estimated elasticity is less than 1. In addition, the elasticity becomes slightly smaller, that is, 0.070, in Column (2) and 0.205 in Column (6). More importantly, the inclusion of establishment dummies further diminishes the elasticity to 0.006 in Column (3) and 0.056 in Column (7), which means that establishment-level TFP only has a minimal impact on establishment-level wage. Furthermore, the first difference estimations in Columns (4) and (8) show much smaller magnitudes.[25]

In sum, the findings suggest that there is no stronger relationship between wage and productivity than theoretically expected, and their positive correlation mainly comes from establishment fixed effects, thus suggesting that the TFP channel of the urban wage premium is quite limited, even if agglomeration economies increase establishment TFP.

Furthermore, establishment-level bonuses are more responsible for establishment-level TFP. As mentioned in Carlsson et al. (2016), individual firm TFP do not affect individual wage in the competitive labor market. This is true in the case of basic pay, but productivity shock at the individual firm level has a modest effect on bonuses.

[Table 5]

7.3 Effects of the Spatial Sorting of Establishments

To clarify the effects of the spatial sorting of establishments, this study investigates how the spatially uneven distribution of establishments affects the urban wage premium. Mion and Naticchioni (2009)

[25] An increase in wage may not immediately follow an increase in productivity, due to short-term wage rigidity. In order to consider the lag structure between wage determination and productivity, one-year lagged productivity is included. The findings do not qualitatively change if one-year or higher lagged productivities are added into the regression.

find that firm size distribution differences affect the urban wage premium. The present study focuses on employment size and financial capital as the variables of establishment size. Figure 5 shows that wages and bonuses become higher as establishment size increases. If large-sized establishments are concentrated in large cities, then the urban wage premium is affected by the spatial sorting of establishments. Indeed, Syverson (2004) indicates that plant size in a high-density market is larger than that in a low-density market. The estimation results also show that the spatial sorting of establishments (in terms of establishment size) leads to an upward effect on the urban wage premium.

Figure 6 illustrates the spatial distribution of establishment mean wages between large and small cities. The solid red (dashed blue) line indicates the establishment mean wage and bonus distributions of cities with above-75 percentile (below-75 percentile) populations. The establishment mean wage in Panels (b) and (d) are additionally controlled for establishment size (i.e., employment size and financial capital) compared to Panels (a) and (c). Clearly, establishment size control reduces the gap of distributions between large and small cities.

Table 6 presents the estimation results of the quantile approach suggested by Combes et al. (2012). Establishment size control respectively decreases the relative shift parameter $\hat{A}$ ($0.0622 \rightarrow 0.0497$) for relative wages and $\hat{A}$ ($0.1486 \rightarrow 0.1146$) for relative bonuses. In other words, establishment size distribution differences, respectively, explains the differences in hourly wage by 1.3% ($\approx \exp(0.013)-1$) and the differences in bonuses by 3.5% ($\approx \exp(0.034) - 1$) between cities with above- and below-75 percentile populations. The results suggest that the urban wage premium is partly explained by the spatial sorting of establishments.

[Table 6 and Figures 5–6]

8 Conclusion

This study has investigated the channels of the urban wage premium using matched employer–employee data in the Japanese manufacturing sector. Although recent studies in the empirics of agglomeration economies have attracted the attentions of numerous researchers, but it is important to carefully interpret the estimation results when identifying the channels of urban wage premium. Even if fundamental assumptions differ considerably between theoretical models, it is possible to obtain similar regression models in empirical studies, which makes the interpretation inseparable between the different models. This study also has discussed how the standard wage regression approach captures not only the TFP channel of the urban wage premium but also other channels (e.g., firm size effects on wages).

The estimation results have shown that the magnitude of the TFP channel of the urban wage premium becomes smaller than those in the existing literature when using matched employer–employee data. This study has found that large-sized establishments, which are generally located in larger cities, pay higher wages, explaining an upward effect on the urban wage premium. Furthermore, by exploiting the features of the Japanese employment system, this study has compared the channels of the urban wage premium between basic pay and bonuses, and has found that the bonuses are more sensitive to city size than the basic pay. Thus, it is suggested that the establishment-level additional benefits from positive TFP change are not distributed to workers as the basic pay in the Japanese labor market, which makes the urban wage premium of hourly wage small and that of bonuses large.

The findings has important implications for future research. Although it is widely believed that agglomeration economies, on average, increase individual wages through firm TFP, the local benefits of this channel might be limited. In other words, it is implied that there are other channels that can explain the urban wage premium. This study also suggests that both the spatial sorting of workers' skills and the spatial sorting of establishments can affect the urban wage premium. As recently indicated by Dauth et al. (2016), better matching between high skilled workers and productive firms in large cities can lead to the urban wage premium. In addition, according to de la Roca and Puga (2017), longer work experience in cities makes workers more productive, which can lead to higher wages. Although this study enriches our understanding of channels of the urban wage premium, further research is necessary to "open the black box" of the urban wage premium.

References

[1] Acemoglu, Daron and Joshua Angrist (2001) "How large are human-capital externalities? Evidence from compulsory-schooling laws," in Bernanke, Ben S. and Kenneth Rogoff eds. *NBER Macroeconomics Annual 2000* Vol. 15: MIT Press, Chap. 1, pp. 9–74.

[2] Andersson, Martin, Johan Klaesson, and Johan P Larsson (2014) "The sources of the urban wage premium by worker skills: Spatial sorting or agglomeration economies?" *Papers in Regional Science* 93(4), pp. 727–747.

[3] Baum-Snow, Nathaniel and Ronni Pavan (2012) "Understanding the city size wage gap," *Review of Economic Studies* 79(1), pp. 88–127.

[4] Brown, Charles and James Medoff (1989) "The employer size-wage effect," *Journal of Political*

Economy 97(5), pp. 1027–1059.

[5] Carlsson, Mikael, Julián Messina, and Oskar Nordström Skans (2016) "Wage adjustment and productivity shocks," *Economic Journal* 126(595), pp. 1739–1773.

[6] Ciccone, Antonio and Robert E. Hall (1996) "Productivity and the density of economic activity," *American Economic Review* 86(1), pp. 54–70.

[7] Combes, Pierre-Philippe and Laurent Gobillon (2015) "The empirics of agglomeration economies," in Duranton, Gilles, J. Vernon Henderson, and William C. Strange eds. *Handbook of Regional and Urban Economics* Vol. 5, Amsterdam: Elsevier, Chap. 5, pp. 247–348.

[8] Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon (2008) "Spatial wage disparities: Sorting matters!," *Journal of Urban Economics* 63(2), pp. 723–742.

[9] Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux (2010) "Estimating agglomeration economies with history, geology, and worker effects," in Glaeser, Edward L. ed. *Agglomeration Economics*, Chicago: University of Chicago Press, Chap. 1, pp. 15–66.

[10] Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon (2011) "The identification of agglomeration economies," *Journal of Economic Geography* 11(2), pp. 253–266.

[11] Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux (2012) "The productivity advantages of large cities: Distinguishing agglomeration from firm selection," *Econometrica* 80(6), pp. 2543–2594.

[12] Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum (2016) "Spatial wage disparities: Workers, firms and assortative matching." Mimeo.
(URL: `https://sites.google.com/site/jenssuedekum/home/working-papers`).

[13] D'Costa, Sabine and Henry G. Overman (2014) "The urban wage growth premium: Sorting or learning?" *Regional Science and Urban Economics* 48(C), pp. 168–179.

[14] de la Roca, Jorge and Diego Puga (2017) "Learning by working in big cities," *Review of Economic Studies* 84(1), pp. 106–142.

[15] Duranton, Gilles (2016) "Agglomeration effects in Colombia," *Journal of Regional Science* 56(2), pp. 210–238.

[16] Glaeser, Edward L. and David C. Maré (2001) "Cities and skills," *Journal of Labor Economics* 19(2), pp. 316–342.

[17] Glaeser, Edward L. and Matthew G. Resseger (2010) "The complementarity between cities and skills," *Journal of Regional Science* 50(1), pp. 221–244.

[18] Gould, E. D. (2007) "Cities, workers, and wages: A structural analysis of the urban wage

premium," *Review of Economic Studies* 74(2), pp. 477–506.

[19] Hashimoto, Masanori (1990) *The Japanese Labor Market in a Comparative Perspective with the United States: A Transaction-Cost Interpretation*, Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

[20] Heuermann, Daniel, Benedikt Halfdanarson, and Jens Süedekum (2010) "Human capital externalities and the urban wage premium: Two literatures and their interrelations," *Urban Studies* 47(4), pp. 749–767.

[21] Ito, Takatoshi (1991) "The Labor Market," in *The Japanese Economy*, Cambridge, MA: MIT Press, Chap. 8, pp. 209–258.

[22] Kato, Takao (2016) "Productivity, wages and unions in Japan." ILO Conditions of Work and Employment Series No. 73.

[23] Kawaguchi, Daiji and Ryo Kambayashi (2010) "Construction and usage of matched data for government statistics," in Kitamura, Yukinobu ed. *Applied Microeconometrics*, Tokyo: Nihon Hyoronsha, Chap. 5, pp. 131–160. (in Japanese).

[24] Kondo, Keisuke (2016) "Testing for agglomeration economies and firm selection in spatial productivity differences: The case of Japan." RIETI Discussion Paper No. 16-E-098.

[25] Levinsohn, James and Amil Petrin (2003) "Estimating production functions using inputs to control for unobservables," *Review of Economic Studies* 70(2), pp. 317–341.

[26] Maré, David and Daniel Graham (2013) "Agglomeration elasticities and firm heterogeneity," *Journal of Urban Economics* 75(C), pp. 44–56.

[27] Matano, Alessia and Paolo Naticchioni (2012) "Wage distribution and the spatial sorting of workers," *Journal of Economic Geography* 12(2), pp. 379–408.

[28] ——— (2016) "What drives the urban wage premium? Evidence along the wage distribution," *Journal of Regional Science* 56(2), pp. 191–209.

[29] Mion, Giordano and Paolo Naticchioni (2009) "The spatial sorting and matching of skills and firms," *Canadian Journal of Economics* 42(1), pp. 28–55.

[30] Moretti, Enrico (2004) "Human capital externalities in cities," in Henderson, J. Vernon and Jacques-François Thisse eds. *Handbook of Regional and Urban Economics* Vol. 4, Amsterdam: Elsevier, Chap. 51, pp. 2243–2291.

[31] Morikawa, Masayuki (2011) "Urban density, human capital, and productivity: An empirical analysis using wage data." RIETI DP 11-E-060.

[32] Petrin, Amil and James Levinsohn (2012) "Measuring aggregate productivity growth using

plant-level data," *RAND Journal of Economics* 43(4), pp. 705–725.

[33] Syverson, Chad (2004) "Market structure and productivity: A concrete example," *Journal of Political Economy* 112(6), pp. 1181–1222.

[34] Vincenty, Thaddeus (1975) "Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations," *Survey Review* 23(176), pp. 88–93.

[35] Wooldridge, Jeffrey M. (2009) "On estimating firm-level production functions using proxy variables to control for unobservables," *Economics Letters* 104(3), pp. 112–114.

[36] Zaiki, Masumi, Takashi Oguchi, Yuichi Kagawa, Akiko Takahashi, Shiro Koike, and Masakazu Yamauchi (2005) "Relationships between distribution of Japanese residential areas and topography." CSIS Discussion Paper No. 68. (in Japanese).

Appendix A Constructing Matched Employer–Employee Data

Following Kawaguchi and Kambayashi (2010), the present study uses the information on the prefecture, municipality, and the establishments' telephone numbers to match the establishments between the Basic Survey of Wage Structure (BSWS) and the Census of Manufacture (CM) (Economic Census for Business Activity only in 2011) from 1993 to 2013.

The sampling design of the BSWS is based on the establishment level and thus individual workers have establishment IDs. The establishment lists obtained by the Establishment and Enterprise Census (EEC) and the Economic Census for Business Frame (ECBF) are used for sampling and we can easily match them with the individual workers' data via the establishment IDs. Therefore, in order to create a matched employer–employee data, we simply need to match any other establishment-level datasets with these establishment lists.

The data construction proceeds as follows. First, the establishment-level panel data are constructed from the CM and the Economic Census for Business Activity (ECBA) from the period 1993–2013. The inter-temporal connection table of establishment IDs is available from the Ministry of Economy, Trade and Industry (METI). Second, the establishment-level panel are created from the EEC and the ECBF. The EEC was conducted in 1991, 1994, 1996, 1999, 2001, 2004, and 2006, and afterwards the ECBF was conducted in 2009. After the 2001 EEC, such data include establishment IDs allocated in the previous censuses when the establishments were surveyed. Third, the establishment-level panel data of the CM are matched with the establishment data of the EEC and the ECBF for each corresponding year, except 1996 and 1999.[26] The panel structure of the establishment-level data enables us to retrospectively allocate the establishment IDs year-by-year between 1991, 1994, 1996, 1999, 2001, 2004, 2006, and 2009.

[Table A.1]

[26]The 1993 Census of Manufacture and the 1991 Establishment and Enterprise Census are exceptionally matched. The telephone numbers are not included in the 1996 and 1999 Establishment and Enterprise Censuses. The 2001 Establishment and Enterprise Census includes the establishment IDs for the 1996 and 1999 Establishment and Enterprise Censuses in 1996 and 1999. Thus, the establishments surveyed in the 1996 and/or 1999 are dropped from the matched data.

Appendix B TFP Estimation

Taking the logarithm of the Cobb-Douglas production function (3), we have

$$\log(q_{jt}) = \xi_0 + \xi_l \log(l_{jt}) + \xi_k \log(k_{jt}) + \omega_{jt} + e_{jt} \tag{12}$$

where q_{jt} is the value added, l_{jt} is the number of workers, and k_{jt} is the capital stock. The error term is assumed to consist of two components: the unobserved productivity ω_{jt}, which affects the establishment's investment decisions and the i.i.d. idiosyncratic shock u_{jt}, which has no impact on the establishment's investment decisions.

Since establishment-level data include no information on workers' skills, the annual hours worked are applied. In addition, production function is estimated using the Wooldridge-Levinsohn-Petrin approach (Levinsohn and Petrin, 2003; Petrin and Levinsohn, 2012; Wooldridge, 2009). After obtaining consistent estimates for the labor and capital elasticities of output ($\hat{\xi}_l$ and $\hat{\xi}_k$), we calculate establishment-level TFP as follows:

$$\log(\hat{A}_{jat}) = \log(q_{jt}) - \hat{\xi}_l \log(l_{jt}) - \hat{\xi}_k \log(k_{jt}).$$

In the estimation, the cost of electricity consumed is used as a proxy variable of productivity shock, while the instrumental variable for labor input is its lagged variable. In order to consider heterogeneity in production technology across industries, establishment-level TFP is estimated by a two-digit level industry. Finally, to make TFP comparable across industries, industry fixed effects are removed. Figure B.1 shows the estimated labor and capital elasticities of output by industry.

[Figure B.1]

Appendix C Two-Step Estimation Approach for TFP and Wages

The case of establishment-year mean and area-year mean bonuses, $\log(\bar{b}_{jast})$ and $\log(\bar{b}_{at})$, is analogous to that of hourly wage. Therefore, it is not explained below.

C.1 Establishment-Year and Area-Year Mean Wages

This study computes establishment-level the mean wage after controlling for individual characteristics as follows:

$$\log(w_{ijast}) = \psi_{jt} + \boldsymbol{X}_{it}\boldsymbol{\beta} + u_{ijast},$$
$$\log(\overline{w}_{jast}) = \hat{\psi}_{jt} - \bar{\hat{\psi}}_{jt},$$

where the establishment-year mean wage is calculated as residual wages averaged across workers by establishment. Note that industry control is not performed in this stage.

In the same manner, the mean wage for each area and year from individual workers' data is computed as follows:

$$\log(w_{ijast}) = \eta_{at} + \boldsymbol{X}_{it}\boldsymbol{\beta} + \gamma_s + u_{ijast},$$
$$\log(\overline{w}_{at}^{\text{Total1}}) = \hat{\eta}_{at} - \bar{\hat{\eta}}_{at},$$

This area-year mean wage $\log(\overline{w}_{at}^{\text{Total1}})$ is intended to control for the fact that the quality of workers is not equally distributed in a geographical space, and unobservable individual effects cannot be controlled for.

The estimation results of the first-step regressions using individual workers' data are shown in Table C.1. In Column (1), the estimates of establishment-year dummies are used as the mean wage for each establishment and year at the second step. In Column (2), the estimates of area-year dummies are used as the mean wage for each municipality and year at the second step. Note that the estimates of the dummies are the geometric mean, since the dependent variable is expressed on the log scale.

[Table C.1]

In addition, the mean wage for each area and year controlled for individual characteristics and firm characteristics is computed as follows:

$$\log(\overline{w}_{jast}) = \eta_{at} + \boldsymbol{H}_{jt}\boldsymbol{\varphi} + \gamma_s + u_{jast},$$
$$\log(\overline{w}_{at}^{\text{Type}}) = \hat{\eta}_{at} - \bar{\hat{\eta}}_{at},$$

where Type ∈ {Total2, Total3}, and $\boldsymbol{H}$ is a vector of firm characteristics (TFP, employment size, and capital). The wage $\overline{w}_{at}^{\text{Total2}}$ is estimated using only TFP, while the wage $\overline{w}_{at}^{\text{Total3}}$ is estimated using all three variables of firm characteristics. The estimation results are shown in Table C.2

[Table C.2]

C.2 Area-Year Mean TFP

This study computes the area-year mean TFP after controlling for establishment characteristics as follows:

$$\log(\hat{A}_{jast}) = \eta_{at} + \gamma_s + u_{jast},$$
$$\log(\overline{\text{TFP}}_{at}) = \hat{\eta}_{at} - \bar{\hat{\eta}}_{at}.$$

Note that the estimates of dummies $\hat{\eta}_{at}$ are the geometric mean, since the dependent variable is expressed on the log scale. In addition, they are rescaled as relative values to the mean.

Table 1: Descriptive Statistics, 1993–2013

Variables	Obs.	Mean	S.D.	Median
Individual Characteristics				
Log(Hourly Wage)	3060323	7.369	0.414	7.358
Log(Bonus)	3060323	8.920	0.912	9.075
Log(Relative Hourly Wage to Minimum Wage)	3060323	0.907	0.403	0.893
Log(Relative Bonus to Minimum Wage)	3060323	2.459	0.910	2.615
Log(Prefecture Minimum Hourly Wage)	3060323	6.461	0.076	6.465
D(1=High School)	3060323	0.635	0.481	1.000
D(1=Junior College)	3060323	0.072	0.258	0.000
D(1=University)	3060323	0.163	0.369	0.000
Age	3060323	40.087	11.736	40.000
Working Years	3060323	14.315	10.712	12.000
Dummy (1=White Collar)	3060323	0.370	0.483	0.000
Dummy (1=Non-Regular Worker)	3060323	0.036	0.186	0.000
Dummy (1=Division Manager)	3060323	0.015	0.120	0.000
Dummy (1=Section Chief)	3060323	0.038	0.192	0.000
Dummy (1=Female)	3060323	0.251	0.433	0.000
Establishment Characteristics				
Log(Establishment Mean Relative Hourly Wage)	100388	0.000	0.177	−0.002
Log(Establishment Mean Relative Bonus)	100388	0.000	0.618	0.133
Log(Establishment Level TFP)	100388	0.005	0.855	−0.030
Log(Employment Size)	100388	8.005	1.041	7.805
Log(Financial Capital)	100388	9.837	2.669	9.068
Regional Characteristics				
Log(Area Mean TFP)	21482	−0.045	0.554	−0.037
Log(Area Mean Relative Hourly Wage) for Total1	21482	−0.037	0.128	−0.029
Log(Area Mean Relative Hourly Wage) for Total2	21482	−0.019	0.119	−0.015
Log(Area Mean Relative Hourly Wage) for Total3	21482	−0.016	0.109	−0.012
Log(Area Mean Relative Bonus) for Total1	21482	−0.097	0.437	−0.011
Log(Area Mean Relative Bonus) for Total2	21482	−0.042	0.428	0.025
Log(Area Mean Relative Bonus) for Total3	21482	−0.033	0.407	0.023
Log(Population)	21482	13.852	1.366	13.733
Log(Manufacturing Establishments)	21482	7.812	1.401	7.704
Share of University Graduates	21482	0.127	0.048	0.122
Log(Population Density in 1930)	21482	5.466	0.939	5.453
Log(Mean Altitude)	21482	4.416	1.283	4.585

Note: Hourly wage is deflated by the consumer price index (2010=1). Population is expressed as the local sum of municipal population within the circle of 30 km radius from the centroid of the municipality. Population density in 1930 is expressed in total population per kilometers squared. Mean altitude is expressed in km. The uppermost and lowermost 0.05 percentile of the distribution of individual hourly wage and TFP are excluded from the sample as extreme outliers, respectively. The uppermost and lowermost 0.5 percentile of the distribution of area mean relative hourly wage and bonuses (based on for Total 1) are excluded from the sample as extreme outliers, respectively.

Table 2: Estimation Results for Hourly Wage and City Size

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Population)	0.026***	0.022***	0.016***		0.023***	0.019***	0.013***
	(0.001)	(0.001)	(0.001)		(0.002)	(0.001)	(0.001)
First Stage Estimation							
Log(Population Density 1930)				1.209***			
				(0.026)			
Log(Mean Altitude)				−0.074***			
				(0.015)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.204	0.206	0.207		0.203	0.205	0.207
First Stage F-value					1775.229	1775.229	1775.229
Overidentification (p-value)					0.161	0.215	0.042

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Area-mean hourly wages relative to the prefectural minimum wage are used to control for spatial price difference. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 3: Estimation Results for Bonuses and City Size

	Dependent Variable: $\log(\bar{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Population)	0.059***	0.043***	0.027***		0.049***	0.035***	0.020***
	(0.004)	(0.004)	(0.003)		(0.005)	(0.004)	(0.004)
First Stage Estimation							
Log(Population Density 1930)				1.209***			
				(0.026)			
Log(Mean Altitude)				−0.074***			
				(0.015)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.197	0.201	0.210		0.196	0.200	0.209
First Stage F-value					1775.229	1775.229	1775.229
Overidentification (p-value)					0.287	0.393	0.179

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Area-mean bonuses relative to the prefectural minimum wage are used to control for spatial price difference. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 4: Estimation Results for TFP and City Size

	Dependent Variable: $\log\left(\overline{\text{TFP}}_{at}\right)$		
	OLS	IV	
Explanatory Variables	(1)	(2)	(3)
Log(Population)	0.090***		0.084***
	(0.006)		(0.007)
First Stage Estimation			
Log(Population Density 1930)		1.209***	
		(0.026)	
Log(Mean Altitude)		−0.074***	
		(0.015)	
Regional Block–Year Dummy	Yes	Yes	Yes
Number of Observations	21482	21482	21482
Number of Municipalities	1512	1512	1512
Adjusted R^2	0.070		0.070
First Stage F-value			1775.229
Overidentification (p-value)			0.399

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 5: Wage and TFP Linkages

	Dependent Variable: $\log(\overline{w}_{jast})$				Dependent Variable: $\log(\overline{b}_{jast})$			
	OLS	OLS	FE	FD	OLS	OLS	FE	FD
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log(TFP_{jt})	0.081***	0.070***	0.006***	0.002**	0.237***	0.205***	0.056***	0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)	(0.005)	(0.005)	(0.004)
Area Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Establishment Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Two-Digit Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	100388	100388	100388	45740	45740	45740	45740	21361
Number of Establishment	32169	32169	32169	17567	17567	17567	17567	5017
Adjusted/Within R^2	0.214	0.321	0.261	0.020	0.199	0.293	0.224	0.040

Note: Heteroskedasticity-consistent standard errors clustered at establishment level are in parentheses. Establishment-year mean hourly wage and bonuses relative to the prefectural minimum wage are used to control for spatial price difference. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 6: Spatial Sorting of Establishments

Model	$\hat{A}$	$\hat{D}$	$\hat{S}$	R^2	Obs. (B)	Obs. (A)
	Variable: Relative Hourly Wage to Minimum Wage					
Without Establishment Size Control	0.0622^+	0.9220^+	-0.0002	0.9816	24127	8042
	(0.0022)	(0.0147)	(0.0018)			
With Establishment Size Control	0.0497^+	1.0359^+	0.0002	0.9906	24127	8042
	(0.0021)	(0.0151)	(0.0005)			
	Variable: Relative Bonus to Minimum Wage					
Without Establishment Size Control	0.1486^+	0.8775^+	0.0001	0.9816	24127	8042
	(0.0069)	(0.0162)	(0.0011)			
With Establishment Size Control	0.1146^+	0.8898^+	0.0002	0.9521	24127	8042
	(0.0069)	(0.0179)	(0.0020)			

Note: Bootstrap standard errors are in parentheses, and 100 times of bootstrap sampling with replacement are conducted. The same sample size is used for each bootstrap sampling. + denotes that $\hat{A}$ and $\hat{S}$ are significantly different from 0 at the 5% level, and $\hat{D}$ is significantly different from 1 at the 5% level. Obs. (B) denotes the number of observations for below-75percentile dense cities and Obs. (A) denotes the number of observations for above-75percentile dense cities.

Table A.1: Sample Size of Matched Employer-Employee Data in Manufacturing Sector

Year	Sample for This Study		Matching Results		
	BSWS		CM↔BSWS	BSWS (Full)	Matching Rate
	# Workers	# Estab.	# Estab.	# Estab.	%
1993	171,114	5,325	11,750	16,353	71.9%
1994	158,748	4,830	10,599	15,158	69.9%
1995	165,703	5,077	11,063	16,111	68.7%
1996	206,322	6,737	14,915	19,437	76.7%
1997	201,159	6,622	14,790	19,539	75.7%
1998	182,849	5,908	12,988	18,523	70.1%
1999	181,756	5,824	12,855	18,021	71.3%
2000	173,819	5,578	12,443	17,270	72.0%
2001	168,553	5,419	12,072	16,566	72.9%
2002	159,988	5,239	12,954	16,847	76.9%
2003	156,938	5,161	12,845	16,580	77.5%
2004	197,799	6,862	14,520	17,248	84.2%
2005	101,185	3,467	8,157	9,277	87.9%
2006	117,239	3,916	8,781	10,246	85.7%
2007	111,474	3,822	8,429	9,876	85.3%
2008	113,252	3,893	8,981	10,346	86.8%
2009	109,235	3,707	9,207	10,592	86.9%
2010	104,547	3,517	8,755	9,973	87.8%
2011	97,100	3,223	8,709	9,881	88.1%
2012	93,502	3,239	8,117	10,616	76.5%
2013	88,041	3,022	8,008	10,348	77.4%
Total	3,060,323	100,388	230,938	298,808	77.3%

Note: BSWS stands for Basic Survey of Wage Structure. CM stands for Census of Manufacture.

Table C.1: First-Step Estimation Results for Two-Step Regressions

	Dependent Variable: $\log(w_{ijast})$		Dependent Variable: $\log(b_{ijast})$	
	For Area-Year Mean Wage	For Estab.–Year Mean Wage	For Area-Year Mean Wage	For Estab.–Year Mean Wage
Explanatory Variables	(1)	(2)		
Area $\times$ **Year Dummy** (η_{at})	Yes	No	Yes	No
Establishment $\times$ **Year Dummy** (ψ_{jt})	No	Yes	No	Yes
D(1=High School)	0.072***	0.051***	0.126***	0.073***
	(0.001)	(0.001)	(0.003)	(0.002)
D(1=Junior College)	0.099***	0.078***	0.183***	0.130***
	(0.002)	(0.001)	(0.004)	(0.002)
D(1=University)	0.162***	0.128***	0.300***	0.196***
	(0.002)	(0.001)	(0.005)	(0.002)
Age	0.039***	0.039***	0.030***	0.035***
	(0.000)	(0.000)	(0.001)	(0.000)
Age Squared ($\times$1/100)	−0.044***	−0.041***	−0.044***	−0.040***
	(0.000)	(0.000)	(0.001)	(0.001)
Working Years	0.016***	0.013***	0.068***	0.057***
	(0.000)	(0.000)	(0.000)	(0.000)
Working Years Squared ($\times$1/100)	−0.004***	−0.004***	−0.091***	−0.088***
	(0.000)	(0.000)	(0.001)	(0.001)
D(1=White Collar)	0.083***	0.073***	0.191***	0.171***
	(0.001)	(0.001)	(0.002)	(0.001)
D(1=Non-Regular Worker)	−0.164***	−0.228***	−0.786***	−0.969***
	(0.003)	(0.003)	(0.013)	(0.012)
D(1=Division Manager)	0.340***	0.354***	0.420***	0.424***
	(0.002)	(0.002)	(0.005)	(0.004)
D(1=Section Chief)	0.193***	0.204***	0.249***	0.244***
	(0.002)	(0.001)	(0.004)	(0.002)
D(1=Female)	−0.326***	−0.294***	−0.384***	−0.327***
	(0.001)	(0.001)	(0.003)	(0.002)
Two-Digit Industry Dummies	Yes	Yes	Yes	Yes
Number of Observations	3060323	3060323	3060323	3060323
Adjusted R^2	0.716	0.826	0.567	0.796

Note: Heteroskedasticity-consistent standard errors clustered at the establishment level are in parentheses. Hourly wage and bonuses relative to the prefectural minimum wage are used to control for spatial price difference. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table C.2: Regressions Results for Hourly Wage and Bonus

	Dependent Variable: $\log(\overline{w}_{jast})$		Dependent Variable: $\log(\overline{b}_{jast})$	
	For Area-Year Mean Wage	For Area-Year Mean Wage	For Area-Year Mean Wage	For Area-Year Mean Wage
Explanatory Variables	(1)	(2)	(3)	(4)
Area × **Year Dummy** (η_{at})	Yes	Yes	Yes	Yes
Log(TFP)	0.073***	0.040***	0.224***	0.137***
	(0.001)	(0.001)	(0.004)	(0.004)
Log(Employment Size)		0.011***		0.066***
		(0.001)		(0.004)
Log(Financial Capital)		0.022***		0.046***
		(0.001)		(0.002)
Two-Digit Industry Dummies	Yes	Yes	Yes	Yes
Number of Observations	100388	100388	100388	100388
Adjusted R^2	0.292	0.401	0.268	0.329

Note: Heteroskedasticity-consistent standard errors clustered at the establishment level are in parentheses. Establishment-year mean hourly wage and bonuses relative to the prefectural minimum wage are used to control for spatial price difference. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

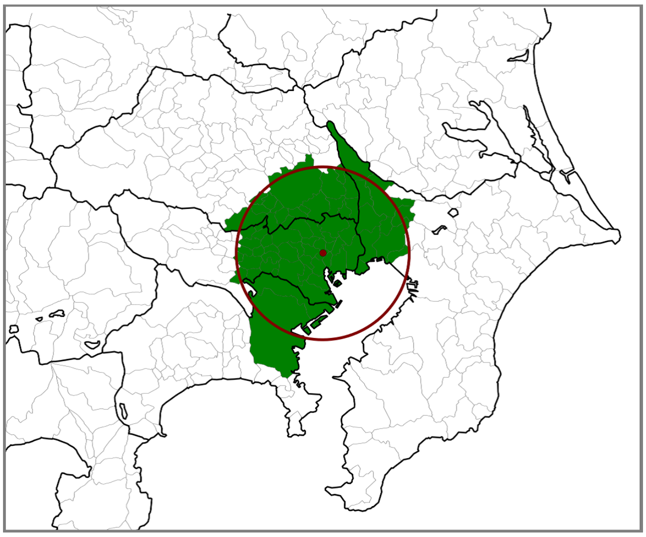
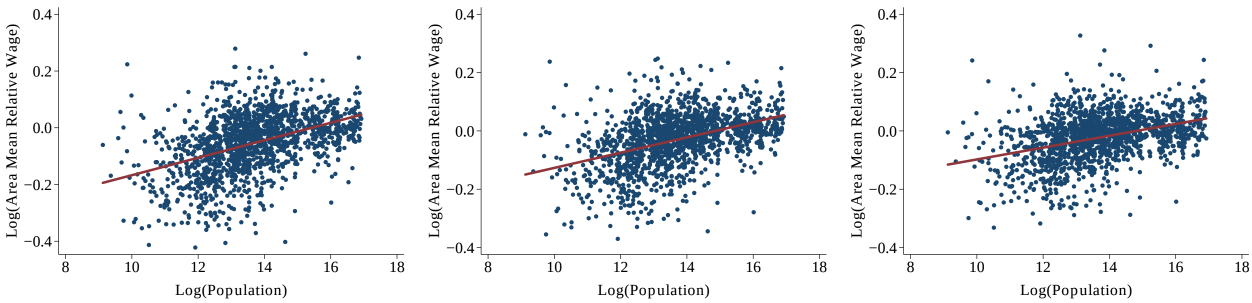


Figure 1: Spatial Unit of Population (Case of Chiyoda-ku,Tokyo)

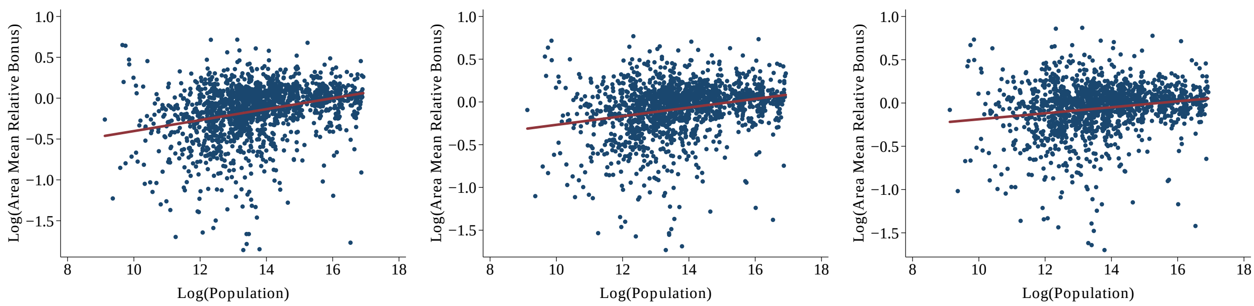
Note: Created by author. Total population within the circle of 30 km radius from the centroid of the municipality is used as a city size. In this case, 68 municipalities are included within the circle of 30 km radius, which means that the population of Chiyoda-ku is calculated as the local sum of municipal population in green colored area.



(a) Relative Hourly Wage 1 (b) Relative Hourly Wage 2 (c) Relative Hourly Wage 3

Figure 2: Hourly Wage and City Size

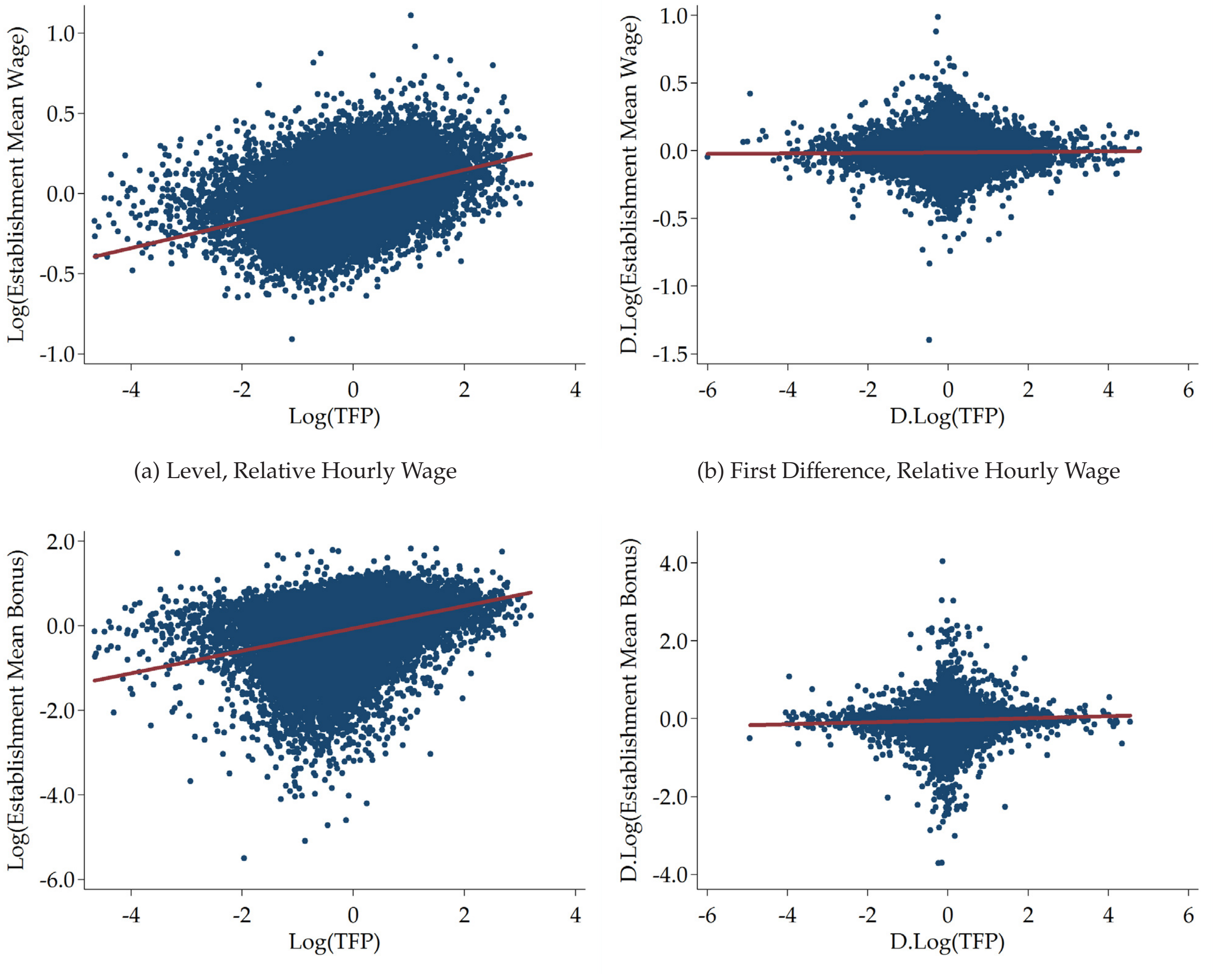
Note: Created by author. Area-mean hourly wages relative to the prefectural minimum wage are used to control for spatial price difference. Area-year mean hourly wages and population are, respectively, averaged across years as $\overline{w}_a^{\text{Type}} = 1/T_a \sum_t \overline{w}_{at}^{\text{Type}}$ and $\overline{\text{Pop}}_a = 1/T_a \sum_t \overline{\text{Pop}}_{at}$, where T_a is the number of years for municipality a observed in the sample.



(a) Relative Bonus 1 (b) Relative Bonus 2 (c) Relative Bonus 3

Figure 3: Bonuses and City Size

Note: Created by author. Area-mean bonuses relative to the prefectural minimum wage are used to control for spatial price difference. Area-year mean bonuses and population are, respectively, averaged across years as $\bar{b}_a^{\text{Type}} = 1/T_a \sum_t \bar{b}_{at}^{\text{Type}}$ and $\overline{\text{Pop}}_a = 1/T_a \sum_t \overline{\text{Pop}}_{at}$, where T_a is the number of years for municipality a observed in the sample.



(a) Level, Relative Hourly Wage

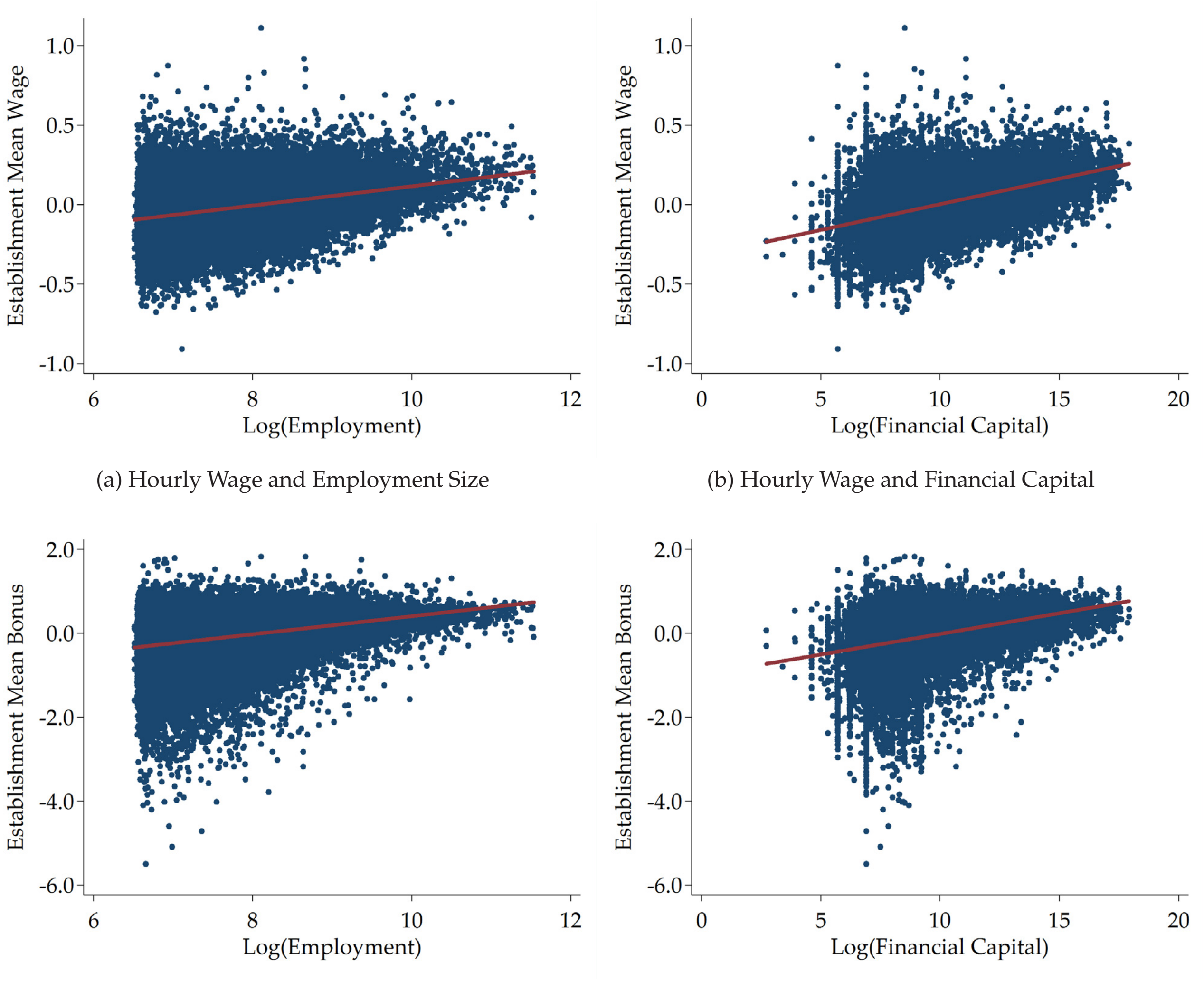
(b) First Difference, Relative Hourly Wage

(c) Level, Relative Bonus

(d) First Difference, Relative Bonus

Figure 4: Establishment Mean Wage and Total Factor Productivity

Note: Created by author. Establishment mean hourly wage and bonuses relative to the prefectural minimum wage are used to control for spatial price difference. In Panel (a), establishment mean wages and TFP are averaged across years as $\overline{w}_j^{\text{Type}} = 1/T_j \sum_t \overline{w}_{jt}^{\text{Type}}$, $\overline{b}_j = 1/T_j \sum_t \overline{b}_{jt}$, and $\overline{\text{TFP}}_j = 1/T_j \sum_t \overline{\text{TFP}}_{jt}$, where T_j is the number of years for establishment j observed in the sample.



(a) Hourly Wage and Employment Size

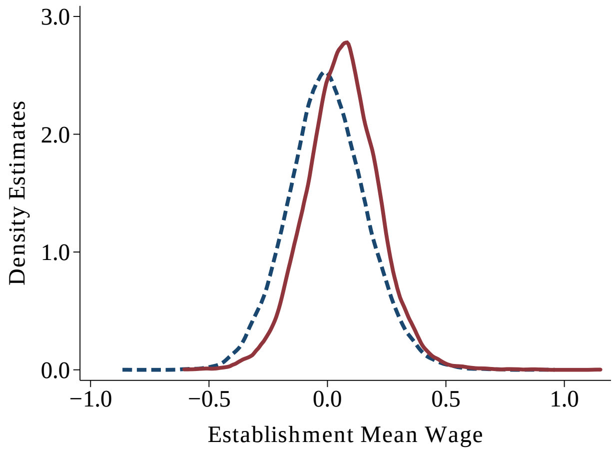
(b) Hourly Wage and Financial Capital

(c) Bonuses and Employment Size

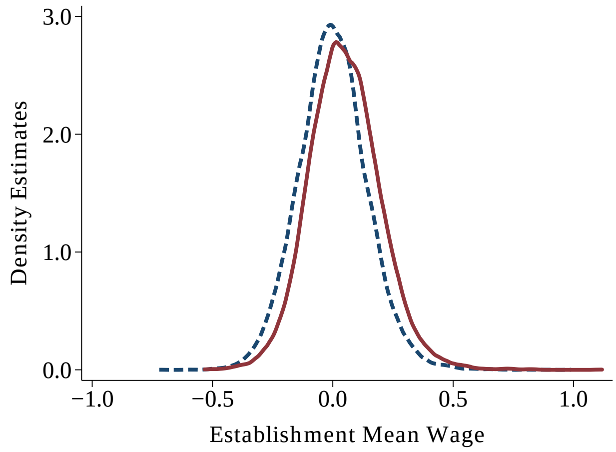
(d) Bonuses and Financial Capital

Figure 5: Establishment Mean Wage and Bonuses and Establishment Size

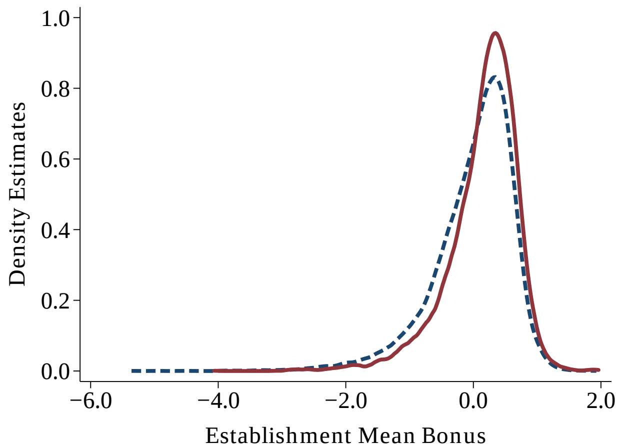
Note: Created by author. Establishment mean wages and bonuses are averaged across years as $\overline{w}_j^{\text{Total1}} = 1/T_j \sum_t \overline{w}_{jt}^{\text{Total1}}$ and $\overline{b}_j^{\text{Total1}} = 1/T_j \sum_t \overline{b}_{jt}^{\text{Total1}}$, where T_j is the number of years for establishment j observed in the sample. The two variables on employment size and financial capital, which express establishment size, are also calculated in the same way.



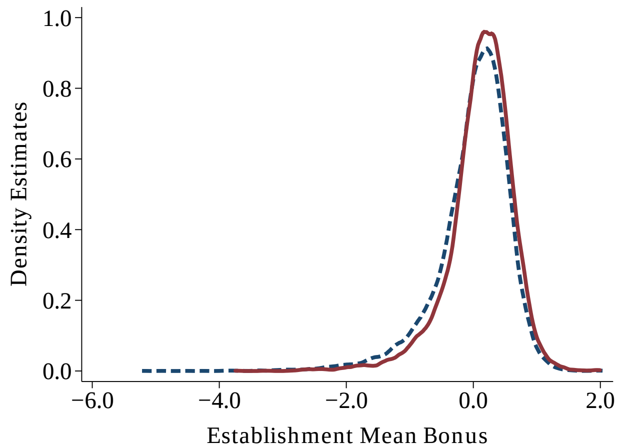
(a) Relative Wages Without Establishment Size Control



(b) Relative Wages With Establishment Size Control



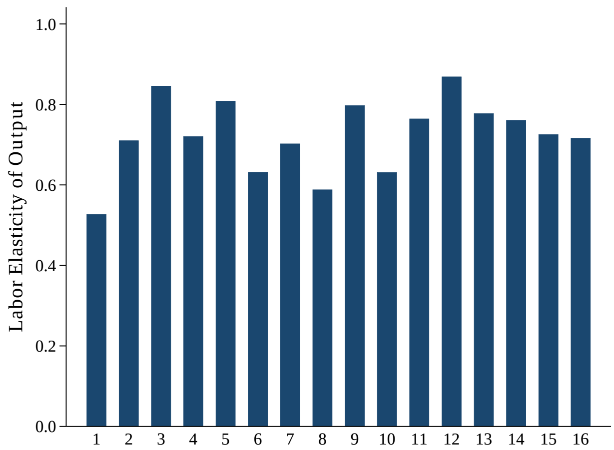
(c) Relative Bonuses Without Establishment Size Control



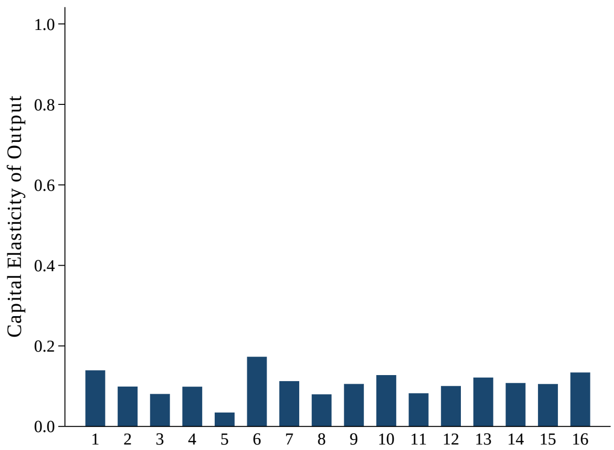
(d) Relative Bonuses With Establishment Size Control

Figure 6: Spatial Distributions of Establishment Mean Wage and Bonus

Note: Created by author. Establishment mean wages and bonuses are averaged across years as $\overline{w}_j^{\text{Total1}} = 1/T_j \sum_t \overline{w}_{jt}^{\text{Total1}}$ and $\overline{b}_j^{\text{Total1}} = 1/T_j \sum_t \overline{b}_{jt}^{\text{Total1}}$, where T_j is the number of years for establishment j observed in the sample. Individual characteristics of workers in establishments are controlled for in all panels. Establishment size (employment size and financial capital) is additionally controlled for in Panels (b) and (d), compared to Panels (a) and (c). The red solid (blue dashed) line indicates the establishment mean wage and bonus distributions of cities with above-75 percentile (below-75 percentile) population.



(a) Labor Elasticity of Output



(b) Capital Elasticity of Output

Figure B.1: Labor and Capital Elasticities of Output to by Sector

Note: Created by author. Numbers in figure indicate industry codes. 1. Food, beverages, tobacco, feed; 2. Textile mill products, leather tanning, leather products, and fur skins; 3. Lumber, wood products, furniture, and fixtures; 4. Pulp, paper and paper products; 5. Printing and allied industries; 6. Chemical and allied products; 7. Plastic products and rubber products; 8. Ceramic, stone and clay products; 9. Iron and steel; 10. Non-ferrous metals and products; 11. Fabricated metal products; 12. General-purpose machinery; 13. Business oriented machinery; 14. Electrical machinery, equipment and supplies, electronic parts, devices and electronic circuits; Information and communication electronics equipment; 15. Transportation equipment; 16. Miscellaneous manufacturing industries

Online Appendix for

Urban Wage Premium Revisited: Evidence from Japanese Matched Employer–Employee Data

Keisuke Kondo*

March 29, 2017

This online appendix provides additional estimation results.

1 Other Local Variables: Share of University Graduates and Number of Manufacturing Establishments

Table O.1–O.4 and Figure O.1–O.2.

2 TFP and City Size

Table O.5–O.6 and Figure O.3.

3 The Cases of Nominal Wages and Bonuses

Table O.7–O.12.

4 Relative Wages and Bonuses Controlled for Dynamic Benefits from Agglomeration Economies

Table O.13–O.23 and Figure O.4–O.5.

*Research Institute of Economy, Trade and Industry (RIETI). 1-3-1 Kasumigaseki, Chiyoda-ku, Tokyo, 100–8901, Japan. (e-mail: kondo-keisuke@rieti.go.jp).

Wages and bonuses are controlled for dynamic benefits from agglomeration economies as follows:

$$\begin{aligned}\log(w_{ijast}) = {} & \theta \log(\text{Pop}_{at}) \\ & + \delta_1\Big(\text{Expr}_{ijat} \times \log(\text{Pop}_{at})\Big) + \delta_2\Big(\text{Expr}^2_{ijat} \times \log(\text{Pop}_{at})\Big) + \delta_3\Big(\text{Expr}^3_{ijat} \times \log(\text{Pop}_{at})\Big) \\ & + \boldsymbol{X}_{it}\boldsymbol{\beta} + \gamma_s + \pi_t + u_{ijast},\end{aligned} \tag{1}$$

where Expr_{ijat} denotes years of working for establishment j in area a of worker i, and $\boldsymbol{X}_{ia}$ is the vector of individual characteristics of worker i (age, gender, years of schooling, work experience, and a dummy variable for non-regular workers). The three cross terms in the second line on the right hand side measure dynamic skill upgrading, which depends on the size of city where the employee works.

5 Nominal Wages and Bonuses Controlled for Dynamic Benefits from Agglomeration Economies

Table O.24–O.31.

Table O.1: Estimation Results for Relative Hourly Wage and Share of University Graduates

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of University Graduates	0.731***	0.609***	0.436***		0.768***	0.632***	0.450***
	(0.045)	(0.039)	(0.036)		(0.055)	(0.048)	(0.044)
First Stage Estimation							
Log(Population Density 1930)				0.038***			
				(0.001)			
Log(Mean Altitude)				0.000			
				(0.001)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.197	0.201	0.204		0.197	0.201	0.204
First Stage F-value					1042.386	1042.386	1042.386
Overidentification (p-value)					0.732	0.843	0.248

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.2: Estimation Results for Relative Bonuses and Share of University Graduates

	Dependent Variable: $\log(\bar{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of University Graduates	1.716***	1.286***	0.839***		1.641***	1.152***	0.672***
	(0.124)	(0.111)	(0.101)		(0.158)	(0.139)	(0.127)
First Stage Estimation							
Log(Population Density 1930)				0.038***			
				(0.001)			
Log(Mean Altitude)				0.000			
				(0.001)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.196	0.201	0.211		0.196	0.201	0.210
First Stage F-value					1042.386	1042.386	1042.386
Overidentification (p-value)					0.763	0.800	0.340

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.3: Estimation Results for Relative Hourly Wage and Manufacturing Establishments

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Establishments)	0.028***	0.023***	0.017***		0.024***	0.020***	0.014***
	(0.001)	(0.001)	(0.001)		(0.002)	(0.001)	(0.001)
First Stage Estimation							
Log(Population Density 1930)				1.226***			
				(0.025)			
Log(Mean Altitude)				−0.013			
				(0.017)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.216	0.213	0.213		0.215	0.212	0.212
First Stage F-value					1557.409	1557.409	1557.409
Overidentification (p-value)					0.506	0.609	0.153

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.4: Estimation Results for Relative Bonuses and Manufacturing Establishments

	Dependent Variable: $\log(\bar{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Establishments)	0.064***	0.046***	0.030***		0.051***	0.036***	0.021***
	(0.004)	(0.004)	(0.003)		(0.005)	(0.004)	(0.004)
First Stage Estimation							
Log(Population Density 1930)				1.226***			
				(0.025)			
Log(Mean Altitude)				−0.013			
				(0.017)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.204	0.204	0.212		0.203	0.203	0.211
First Stage F-value					1557.409	1557.409	1557.409
Overidentification (p-value)					0.596	0.666	0.285

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.5: Estimation Results for TFP and Share of University Graduates

	Dependent Variable: $\log\left(\overline{\text{TFP}}_{at}\right)$		
	OLS	IV	
Explanatory Variables	(1)	(2)	(3)
Share of University Graduates	2.513***		2.784***
	(0.162)		(0.220)
First Stage Estimation			
Log(Population Density 1930)		0.038***	
		(0.001)	
Log(Mean Altitude)		0.000	
		(0.001)	
Regional Block–Year Dummy	Yes	Yes	Yes
Number of Observations	21482	21482	21482
Number of Municipalities	1512	1512	1512
Adjusted R^2	0.065		0.065
First Stage F-value			1042.386
Overidentification (p-value)			0.899

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.6: Estimation Results for TFP and Manufacturing Establishments

	Dependent Variable: $\log\left(\overline{\text{TFP}}_{at}\right)$		
	OLS	IV	
Explanatory Variables	(1)	(2)	(3)
Log(Establishments)	0.097***		0.086***
	(0.006)		(0.007)
First Stage Estimation			
Log(Population Density 1930)		1.226***	
		(0.025)	
Log(Mean Altitude)		−0.013	
		(0.017)	
Regional Block–Year Dummy	Yes	Yes	Yes
Number of Observations	21482	21482	21482
Number of Municipalities	1512	1512	1512
Adjusted R^2	0.079		0.079
First Stage F-value			1557.409
Overidentification (p-value)			0.873

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.7: Estimation Results for Nominal Wage and City Size

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Population)	0.053***	0.048***	0.042***		0.051***	0.046***	0.041***
	(0.001)	(0.001)	(0.001)		(0.002)	(0.002)	(0.001)
First Stage Estimation							
Log(Population Density 1930)				1.209***			
				(0.026)			
Log(Mean Altitude)				−0.074***			
				(0.015)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.277	0.275	0.259		0.277	0.275	0.259
First Stage F-value					1775.229	1775.229	1775.229
Overidentification (p-value)					0.041	0.067	0.008

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.8: Estimation Results for Nominal Bonuses and City Size

	Dependent Variable: $\log(\bar{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Population)	0.085***	0.069***	0.053***		0.077***	0.062***	0.047***
	(0.004)	(0.004)	(0.004)		(0.005)	(0.004)	(0.004)
First Stage Estimation							
Log(Population Density 1930)				1.209***			
				(0.026)			
Log(Mean Altitude)				−0.074***			
				(0.015)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.152	0.146	0.141		0.152	0.146	0.141
First Stage F-value					1775.229	1775.229	1775.229
Overidentification (p-value)					0.188	0.278	0.111

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.9: Estimation Results for Nominal Wage and Share of University Graduates

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of University Graduates	1.543***	1.404***	1.235***		1.678***	1.532***	1.354***
	(0.046)	(0.040)	(0.037)		(0.061)	(0.054)	(0.049)
First Stage Estimation							
Log(Population Density 1930)				0.038***			
				(0.001)			
Log(Mean Altitude)				0.000			
				(0.001)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.275	0.274	0.261		0.273	0.272	0.259
First Stage F-value					1042.386	1042.386	1042.386
Overidentification (p-value)					0.947	0.664	0.778

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.10: Estimation Results for Nominal Bonuses and Share of University Graduates

	Dependent Variable: $\log(\bar{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of University Graduates	2.528***	2.082***	1.638***		2.552***	2.051***	1.575***
	(0.127)	(0.113)	(0.103)		(0.165)	(0.145)	(0.133)
First Stage Estimation							
Log(Population Density 1930)				0.038***			
				(0.001)			
Log(Mean Altitude)				0.000			
				(0.001)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.153	0.148	0.144		0.153	0.148	0.144
First Stage F-value					1042.386	1042.386	1042.386
Overidentification (p-value)					0.876	0.974	0.490

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.11: Estimation Results for Nominal Wage and Manufacturing Establishments

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Establishments)	0.055***	0.050***	0.044***		0.052***	0.047***	0.042***
	(0.001)	(0.001)	(0.001)		(0.002)	(0.002)	(0.001)
First Stage Estimation							
Log(Population Density 1930)				1.226***			
				(0.025)			
Log(Mean Altitude)				−0.013			
				(0.017)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.311	0.302	0.289		0.310	0.302	0.288
First Stage F-value					1557.409	1557.409	1557.409
Overidentification (p-value)					0.561	0.809	0.300

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.12: Estimation Results for Nominal Bonuses and Manufacturing Establishments

	Dependent Variable: $\log(\bar{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Establishments)	0.091***	0.073***	0.058***		0.079***	0.063***	0.048***
	(0.004)	(0.004)	(0.003)		(0.005)	(0.004)	(0.004)
First Stage Estimation							
Log(Population Density 1930)				1.226***			
				(0.025)			
Log(Mean Altitude)				−0.013			
				(0.017)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.166	0.154	0.148		0.164	0.153	0.147
First Stage F-value					1557.409	1557.409	1557.409
Overidentification (p-value)					0.612	0.728	0.335

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.13: First-Step Estimation Results of Relative Wage for Two-Step Regressions with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(w_{ijast})$			
			For Area-Year Mean Wage	For Estab.–Year Mean Wage
Explanatory Variables	(1)	(2)	(3)	(4)
Area × Year Dummy (η_{at})	No	No	Yes	No
Establishment × Year Dummy (ψ_{jt})	No	No	No	Yes
Log(Population)	0.024***	0.016***		
	(0.001)	(0.001)		
Working Years × Log(Population)		0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)
Working Years Squared × Log(Population)		0.001**	0.001***	0.000
		(0.000)	(0.000)	(0.000)
Working Years Cubed × Log(Population)		−0.009***	−0.007***	−0.006***
		(0.000)	(0.000)	(0.000)
D(1=High School)	0.051***	0.050***	0.071***	0.050***
	(0.002)	(0.002)	(0.001)	(0.001)
D(1=Junior College)	0.069***	0.069***	0.098***	0.078***
	(0.002)	(0.002)	(0.002)	(0.001)
D(1=University)	0.134***	0.133***	0.161***	0.128***
	(0.002)	(0.002)	(0.002)	(0.001)
Age	0.037***	0.037***	0.039***	0.039***
	(0.000)	(0.000)	(0.000)	(0.000)
Age Squared (×1/100)	−0.043***	−0.043***	−0.044***	−0.041***
	(0.000)	(0.000)	(0.000)	(0.000)
Working Years	0.017***	−0.005***	−0.004***	−0.011***
	(0.000)	(0.002)	(0.001)	(0.001)
Working Years Squared (×1/100)	−0.005***	0.051***	0.040***	0.038***
	(0.000)	(0.005)	(0.004)	(0.003)
D(1=White Collar)	0.094***	0.094***	0.084***	0.074***
	(0.001)	(0.001)	(0.001)	(0.001)
D(1=Non-Regular Worker)	−0.192***	−0.192***	−0.165***	−0.229***
	(0.003)	(0.003)	(0.003)	(0.003)
D(1=Division Manager)	0.340***	0.332***	0.333***	0.347***
	(0.003)	(0.003)	(0.003)	(0.002)
D(1=Section Chief)	0.193***	0.187***	0.188***	0.199***
	(0.002)	(0.002)	(0.002)	(0.001)
D(1=Female)	−0.338***	−0.338***	−0.325***	−0.294***
	(0.002)	(0.002)	(0.001)	(0.001)
Two-Digit Industry Dummies	Yes	Yes	Yes	Yes
Number of Observations	3060323	3060323	3060323	3060323
Adjusted R^2	0.656	0.658	0.718	0.827

Note: Heteroskedasticity-consistent standard errors clustered at the establishment level are in parentheses. Hourly wage and bonuses relative to the prefecture minimum wage are used to control for spatial price difference. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.14: First-Step Estimation Results of Relative Bonuses for Two-Step Regressions with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(b_{ijast})$			
			For Area-Year Mean Wage	For Estab.–Year Mean Wage
Explanatory Variables	(1)	(2)	(3)	(4)
Area × Year Dummy (η_{at})	No	No	Yes	No
Establishment × Year Dummy (ψ_{jt})	No	No	No	Yes
Log(Population)	0.062***	0.042***		
	(0.003)	(0.004)		
Working Years × Log(Population)		0.005***	0.005***	0.005***
		(0.000)	(0.000)	(0.000)
Working Years Squared × Log(Population)		−0.025***	−0.025***	−0.027***
		(0.001)	(0.001)	(0.001)
Working Years Cubed × Log(Population)		0.033***	0.039***	0.042***
		(0.001)	(0.001)	(0.000)
D(1=High School)	0.032***	0.036***	0.133***	0.081***
	(0.004)	(0.004)	(0.003)	(0.002)
D(1=Junior College)	0.054***	0.057***	0.190***	0.138***
	(0.006)	(0.006)	(0.004)	(0.002)
D(1=University)	0.186***	0.189***	0.309***	0.206***
	(0.006)	(0.006)	(0.005)	(0.002)
Age	0.024***	0.023***	0.029***	0.034***
	(0.001)	(0.001)	(0.001)	(0.000)
Age Squared (×1/100)	−0.041***	−0.040***	−0.043***	−0.039***
	(0.001)	(0.001)	(0.001)	(0.001)
Working Years	0.070***	0.041***	0.052***	0.036***
	(0.001)	(0.005)	(0.004)	(0.003)
Working Years Squared (×1/100)	−0.093***	−0.012	−0.052***	−0.057***
	(0.001)	(0.014)	(0.011)	(0.007)
D(1=White Collar)	0.222***	0.221***	0.190***	0.170***
	(0.003)	(0.003)	(0.002)	(0.001)
D(1=Non-Regular Worker)	−0.880***	−0.881***	−0.785***	−0.972***
	(0.015)	(0.015)	(0.013)	(0.012)
D(1=Division Manager)	0.441***	0.470***	0.454***	0.462***
	(0.007)	(0.007)	(0.005)	(0.004)
D(1=Section Chief)	0.268***	0.288***	0.273***	0.271***
	(0.005)	(0.005)	(0.004)	(0.002)
D(1=Female)	−0.401***	−0.401***	−0.385***	−0.327***
	(0.004)	(0.004)	(0.003)	(0.002)
Two-Digit Industry Dummies	Yes	Yes	Yes	Yes
Number of Observations	3060323	3060323	3060323	3060323
Adjusted R^2	0.443	0.448	0.574	0.803

Note: Heteroskedasticity-consistent standard errors clustered at the establishment level are in parentheses. Hourly wage and bonuses relative to the prefecture minimum wage are used to control for spatial price difference. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.15: Regressions Results for Relative Wages and Bonuses Controlled for Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\overline{w}_{jast})$		Dependent Variable: $\log(\overline{b}_{jast})$	
	For Area-Year Mean Wage	For Area-Year Mean Wage	For Area-Year Mean Wage	For Area-Year Mean Wage
Explanatory Variables	(1)	(2)	(3)	(4)
Area × Year Dummy (η_{at})	Yes	Yes	Yes	Yes
Log(TFP)	0.072***	0.040***	0.227***	0.139***
	(0.001)	(0.001)	(0.004)	(0.004)
Log(Employment Size)		0.011***		0.066***
		(0.001)		(0.004)
Log(Financial Capital)		0.022***		0.047***
		(0.001)		(0.002)
Two-Digit Industry Dummies	Yes	Yes	Yes	Yes
Number of Observations	100388	100388	100388	100388
Adjusted R^2	0.270	0.382	0.264	0.327

Note: Heteroskedasticity-consistent standard errors clustered at the establishment level are in parentheses. Establishment-year mean hourly wage and bonuses relative to the prefecture minimum wage are used to control for spatial price difference. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.16: Estimation Results for Relative Wage and City Size with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Population)	0.018***	0.010***	0.004***		0.015***	0.007***	0.001
	(0.001)	(0.001)	(0.001)		(0.002)	(0.001)	(0.001)
First Stage Estimation							
Log(Population Density 1930)				1.209***			
				(0.026)			
Log(Mean Altitude)				−0.074***			
				(0.015)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.166	0.158	0.170		0.165	0.157	0.169
First Stage F-value					1775.229	1775.229	1775.229
Overidentification (p-value)					0.159	0.222	0.045

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.17: Estimation Results for Relative Bonuses and City Size with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\bar{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Population)	0.042***	0.020***	0.003		0.032***	0.011**	−0.004
	(0.004)	(0.004)	(0.003)		(0.005)	(0.004)	(0.004)
First Stage Estimation							
Log(Population Density 1930)				1.209***			
				(0.026)			
Log(Mean Altitude)				−0.074***			
				(0.015)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.188	0.192	0.208		0.187	0.192	0.207
First Stage F-value					1775.229	1775.229	1775.229
Overidentification (p-value)					0.273	0.378	0.166

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.18: Estimation Results for Relative Hourly Wage and Share of University Graduates with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of University Graduates	0.512***	0.292***	0.121***		0.495***	0.233***	0.054
	(0.044)	(0.039)	(0.035)		(0.054)	(0.047)	(0.043)
First Stage Estimation							
Log(Population Density 1930)				0.038***			
				(0.001)			
Log(Mean Altitude)				0.000			
				(0.001)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.164	0.159	0.171		0.164	0.158	0.170
First Stage F-value					1042.386	1042.386	1042.386
Overidentification (p-value)					0.467	0.398	0.056

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.19: Estimation Results for Relative Bonuses and Share of University Graduates with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\bar{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of University Graduates	1.267***	0.666***	0.213**		1.073***	0.366***	−0.122
	(0.124)	(0.112)	(0.101)		(0.158)	(0.139)	(0.128)
First Stage Estimation							
Log(Population Density 1930)				0.038***			
				(0.001)			
Log(Mean Altitude)				0.000			
				(0.001)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.189	0.194	0.208		0.188	0.193	0.207
First Stage F-value					1042.386	1042.386	1042.386
Overidentification (p-value)					0.546	0.486	0.145

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.20: Estimation Results for Relative Hourly Wage and Manufacturing Establishments with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Establishments)	0.020***	0.011***	0.005***		0.015***	0.007***	0.002
	(0.001)	(0.001)	(0.001)		(0.002)	(0.001)	(0.001)
First Stage Estimation							
Log(Population Density 1930)				1.226***			
				(0.025)			
Log(Mean Altitude)				−0.013			
				(0.017)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.176	0.163	0.172		0.173	0.161	0.170
First Stage F-value					1557.409	1557.409	1557.409
Overidentification (p-value)					0.349	0.337	0.052

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.21: Estimation Results for Relative Bonuses and Manufacturing Establishments with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\bar{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
		OLS				IV	
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Establishments)	0.048***	0.024***	0.008**		0.033***	0.011***	−0.004
	(0.004)	(0.004)	(0.003)		(0.005)	(0.004)	(0.004)
First Stage Estimation							
Log(Population Density 1930)				1.226***			
				(0.025)			
Log(Mean Altitude)				−0.013			
				(0.017)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.194	0.195	0.209		0.192	0.193	0.207
First Stage F-value					1557.409	1557.409	1557.409
Overidentification (p-value)					0.453	0.453	0.151

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.22: Wages Controlled for Dynamic Benefits from Agglomeration Economies and TFP Linkages

	Dependent Variable: $\log(\overline{w}_{jast})$				Dependent Variable: $\log(\overline{b}_{jast})$			
	OLS	OLS	FE	FD	OLS	OLS	FE	FD
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log(TFP$_{jt}$)	0.079***	0.070***	0.006***	0.002**	0.235***	0.208***	0.058***	0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)	(0.005)	(0.004)	(0.004)
Area Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Establishment Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Two-Digit Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	100388	100388	100388	45740	45740	45740	45740	21361
Number of Establishment	32169	32169	32169	17567	17567	17567	17567	5017
Adjusted/Within R^2	0.205	0.299	0.255	0.020	0.199	0.288	0.230	0.041

Note: Heteroskedasticity-consistent standard errors clustered at establishment level are in parentheses. Establishment-year mean hourly wage and bonuses relative to the prefecture minimum wage are used to control for spatial price difference. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.23: Spatial Sorting of Establishments

Model	$\hat{A}$	$\hat{D}$	$\hat{S}$	R^2	Obs. (B)	Obs. (A)
	Variable: Relative Hourly Wage to Minimum Wage					
Without Establishment Size Control	0.0324^{+}	0.9286^{+}	-0.0001	0.9355	24127	8042
	(0.0023)	(0.0138)	(0.0016)			
With Establishment Size Control	0.0204^{+}	1.0456^{+}	0.0001	0.9563	24127	8042
	(0.0020)	(0.0123)	(0.0006)			
	Variable: Relative Bonuses to Minimum Wage					
Without Establishment Size Control	0.0913^{+}	0.8811^{+}	0.0001	0.9667	24127	8042
	(0.0069)	(0.0145)	(0.0007)			
With Establishment Size Control	0.0573^{+}	0.8935^{+}	0.0001	0.9010	24127	8042
	(0.0132)	(0.0359)	(0.0057)			

Note: Bootstrap standard errors are in parentheses, and 100 times of bootstrap sampling with replacement are conducted. The same sample size is used for each bootstrap sampling. + denotes that $\hat{A}$ and $\hat{S}$ are significantly different from 0 at the 5% level, and $\hat{D}$ is significantly different from 1 at the 5% level. Obs. (B) denotes the number of observations for below-75percentile dense cities and Obs. (A) denotes the number of observations for above-75percentile dense cities.

Table O.24: First-Step Estimation Results of Nominal Wage for Two-Step Regressions with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: log(w_{ijast})			
			For Area-Year Mean Wage	For Estab.–Year Mean Wage
Explanatory Variables	(1)	(2)	(3)	(4)
Area × **Year Dummy** (η_{at})	No	No	Yes	No
Establishment × **Year Dummy** (ψ_{jt})	No	No	No	Yes
Log(Population)	0.049***	0.043***		
	(0.001)	(0.001)		
Working Years × Log(Population)		0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)
Working Years Squared × Log(Population)		−0.000	0.001***	0.000
		(0.000)	(0.000)	(0.000)
Working Years Cubed × Log(Population)		−0.007***	−0.007***	−0.006***
		(0.000)	(0.000)	(0.000)
D(1=High School)	0.082***	0.081***	0.071***	0.050***
	(0.002)	(0.002)	(0.001)	(0.001)
D(1=Junior College)	0.117***	0.116***	0.098***	0.078***
	(0.002)	(0.002)	(0.002)	(0.001)
D(1=University)	0.184***	0.182***	0.161***	0.128***
	(0.002)	(0.002)	(0.002)	(0.001)
Age	0.037***	0.037***	0.039***	0.039***
	(0.000)	(0.000)	(0.000)	(0.000)
Age Squared (×1/100)	−0.043***	−0.043***	−0.044***	−0.041***
	(0.000)	(0.000)	(0.000)	(0.000)
Working Years	0.017***	−0.004**	−0.004***	−0.011***
	(0.000)	(0.002)	(0.001)	(0.001)
Working Years Squared (×1/100)	−0.003***	0.055***	0.040***	0.038***
	(0.000)	(0.005)	(0.004)	(0.003)
D(1=White Collar)	0.087***	0.088***	0.084***	0.074***
	(0.001)	(0.001)	(0.001)	(0.001)
D(1=Non-Regular Worker)	−0.139***	−0.140***	−0.165***	−0.229***
	(0.003)	(0.003)	(0.003)	(0.003)
D(1=Division Manager)	0.329***	0.323***	0.333***	0.347***
	(0.003)	(0.003)	(0.003)	(0.002)
D(1=Section Chief)	0.187***	0.182***	0.188***	0.199***
	(0.002)	(0.002)	(0.002)	(0.001)
D(1=Female)	−0.344***	−0.344***	−0.325***	−0.294***
	(0.002)	(0.002)	(0.001)	(0.001)
Two-Digit Industry Dummies	Yes	Yes	Yes	Yes
Number of Observations	3060323	3060323	3060323	3060323
Adjusted R^2	0.679	0.680	0.733	0.836

Note: Heteroskedasticity-consistent standard errors clustered at the establishment level are in parentheses. Hourly wage and bonuses relative to the prefecture minimum wage are used to control for spatial price difference. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.25: First-Step Estimation Results of Nominal Bonuses for Two-Step Regressions with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(b_{ijast})$			
			For Area-Year Mean Wage	For Estab.–Year Mean Wage
Explanatory Variables	(1)	(2)	(3)	(4)
Area × Year Dummy (η_{at})	No	No	Yes	No
Establishment × Year Dummy (ψ_{jt})	No	No	No	Yes
Log(Population)	0.087***	0.069***		
	(0.003)	(0.004)		
Working Years × Log(Population)		0.005***	0.005***	0.005***
		(0.000)	(0.000)	(0.000)
Working Years Squared × Log(Population)		−0.026***	−0.025***	−0.027***
		(0.001)	(0.001)	(0.001)
Working Years Cubed × Log(Population)		0.034***	0.039***	0.042***
		(0.001)	(0.001)	(0.000)
D(1=High School)	0.063***	0.067***	0.133***	0.081***
	(0.004)	(0.004)	(0.003)	(0.002)
D(1=Junior College)	0.102***	0.105***	0.190***	0.138***
	(0.006)	(0.006)	(0.004)	(0.002)
D(1=University)	0.235***	0.238***	0.309***	0.206***
	(0.006)	(0.006)	(0.005)	(0.002)
Age	0.024***	0.023***	0.029***	0.034***
	(0.001)	(0.001)	(0.001)	(0.000)
Age Squared (×1/100)	−0.041***	−0.040***	−0.043***	−0.039***
	(0.001)	(0.001)	(0.001)	(0.001)
Working Years	0.070***	0.042***	0.052***	0.036***
	(0.001)	(0.005)	(0.004)	(0.003)
Working Years Squared (×1/100)	−0.092***	−0.008	−0.052***	−0.057***
	(0.001)	(0.014)	(0.011)	(0.007)
D(1=White Collar)	0.216***	0.215***	0.190***	0.170***
	(0.003)	(0.003)	(0.002)	(0.001)
D(1=Non-Regular Worker)	−0.828***	−0.829***	−0.785***	−0.972***
	(0.015)	(0.015)	(0.013)	(0.012)
D(1=Division Manager)	0.430***	0.461***	0.454***	0.462***
	(0.007)	(0.007)	(0.005)	(0.004)
D(1=Section Chief)	0.261***	0.283***	0.273***	0.271***
	(0.005)	(0.005)	(0.004)	(0.002)
D(1=Female)	−0.407***	−0.407***	−0.385***	−0.327***
	(0.004)	(0.004)	(0.003)	(0.002)
Two-Digit Industry Dummies	Yes	Yes	Yes	Yes
Number of Observations	3060323	3060323	3060323	3060323
Adjusted R^2	0.459	0.464	0.576	0.804

Note: Heteroskedasticity-consistent standard errors clustered at the establishment level are in parentheses. Hourly wage and bonuses relative to the prefecture minimum wage are used to control for spatial price difference. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.26: Estimation Results for Nominal Wage and City Size with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Population)	0.044***	0.036***	0.030***		0.042***	0.034***	0.029***
	(0.001)	(0.001)	(0.001)		(0.002)	(0.002)	(0.001)
First Stage Estimation							
Log(Population Density 1930)				1.209***			
				(0.026)			
Log(Mean Altitude)				−0.074***			
				(0.015)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.220	0.187	0.165		0.220	0.187	0.165
First Stage					1775.229	1775.229	1775.229
Overidentification					0.040	0.069	0.009

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.27: Estimation Results for Nominal Bonuses and City Size with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\overline{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Population)	0.068***	0.046***	0.029***		0.059***	0.038***	0.023***
	(0.004)	(0.004)	(0.004)		(0.005)	(0.004)	(0.004)
First Stage Estimation							
Log(Population Density 1930)				1.209***			
				(0.026)			
Log(Mean Altitude)				−0.074***			
				(0.015)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.133	0.124	0.125		0.132	0.124	0.124
First Stage					1775.229	1775.229	1775.229
Overidentification					0.178	0.267	0.103

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.28: Estimation Results for Nominal Wage and Share of University Graduates with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of University Graduates	1.324***	1.088***	0.919***		1.405***	1.133***	0.956***
	(0.045)	(0.040)	(0.036)		(0.060)	(0.052)	(0.047)
First Stage Estimation							
Log(Population Density 1930)				0.038***			
				(0.001)			
Log(Mean Altitude)				0.000			
				(0.001)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.226	0.198	0.179		0.226	0.198	0.179
First Stage					1042.386	1042.386	1042.386
Overidentification					0.779	0.886	0.329

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.29: Estimation Results for Nominal Bonuses and Share of University Graduates with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\overline{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of University Graduates	2.079***	1.462***	1.011***		1.984***	1.266***	0.781***
	(0.127)	(0.114)	(0.103)		(0.165)	(0.145)	(0.133)
First Stage Estimation							
Log(Population Density 1930)				0.038***			
				(0.001)			
Log(Mean Altitude)				0.000			
				(0.001)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.137	0.129	0.129		0.137	0.128	0.128
First Stage					1042.386	1042.386	1042.386
Overidentification					0.650	0.636	0.231

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.30: Estimation Results for Nominal Wage and Manufacturing Establishments with Dynamic Benefits from Agglomeration Economies

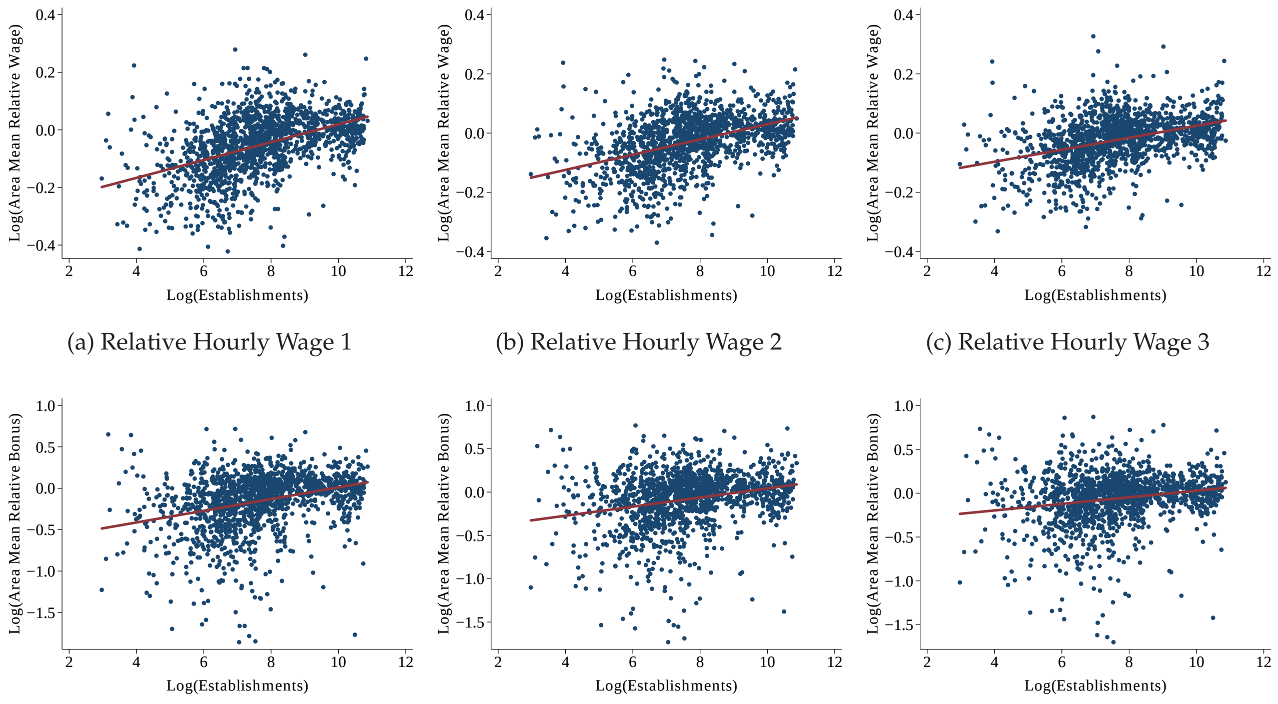
	Dependent Variable: $\log(\overline{w}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Establishments)	0.047***	0.038***	0.032***		0.043***	0.035***	0.029***
	(0.001)	(0.001)	(0.001)		(0.002)	(0.002)	(0.001)
First Stage Estimation							
Log(Population Density 1930)				1.226***			
				(0.025)			
Log(Mean Altitude)				−0.013			
				(0.017)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.255	0.215	0.194		0.253	0.214	0.193
First Stage					1557.409	1557.409	1557.409
Overidentification					0.403	0.507	0.129

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table O.31: Estimation Results for Nominal Bonuses and Manufacturing Establishments with Dynamic Benefits from Agglomeration Economies

	Dependent Variable: $\log(\overline{b}_{at}^{\text{Type}})$						
	Total1	Total2	Total3		Total1	Total2	Total3
	OLS			IV			
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Establishments)	0.075***	0.051***	0.035***		0.061***	0.039***	0.024***
	(0.004)	(0.004)	(0.004)		(0.005)	(0.004)	(0.004)
First Stage Estimation							
Log(Population Density 1930)				1.226***			
				(0.025)			
Log(Mean Altitude)				−0.013			
				(0.017)			
Regional Block–Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wage Control for TFP	No	Yes	Yes		No	Yes	Yes
Wage Control for Establishment Size	No	No	Yes		No	No	Yes
Number of Observations	21482	21482	21482	21482	21482	21482	21482
Number of Municipalities	1512	1512	1512		1512	1512	1512
Adjusted R^2	0.145	0.131	0.130		0.143	0.129	0.128
First Stage					1557.409	1557.409	1557.409
Overidentification					0.469	0.507	0.184

Note: Heteroskedasticity-consistent standard errors clustered at municipal level are in parentheses. Instrumental variables for the logarithm of population are the logarithm of population density in 1930 and logarithm of the mean altitude. Constant is not reported. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

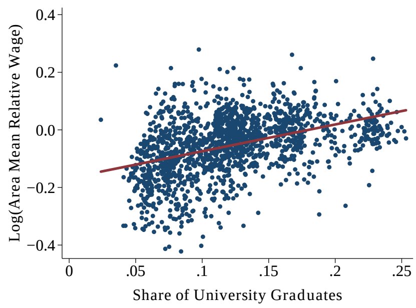


(a) Relative Hourly Wage 1 (b) Relative Hourly Wage 2 (c) Relative Hourly Wage 3

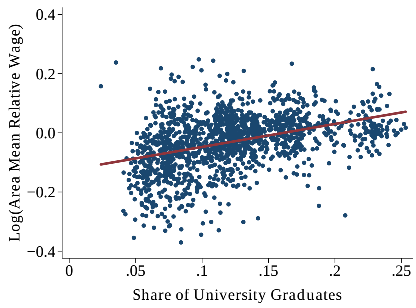
(d) Relative Bonuses 1 (e) Relative Bonuses 2 (f) Relative Bonuses 3

Figure O.1: Wage and Manufacturing Establishments

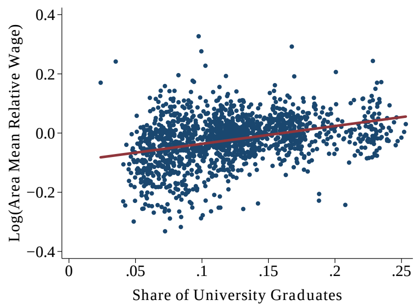
Note: Created by author. Area-year mean wages and manufacturing establishments are, respectively, averaged across years as $\overline{w}_a^{\text{Type}} = 1/T_a \sum_t \overline{w}_{at}^{\text{Type}}$, $\overline{\text{Estb}}_a = 1/T_a \sum_t \overline{\text{Estb}}_{at}$, where T_a is the number of years for municipality a observed in the sample.



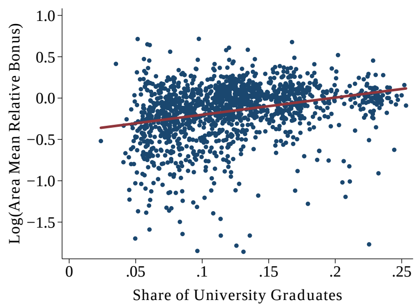
(a) Relative Hourly Wage 1



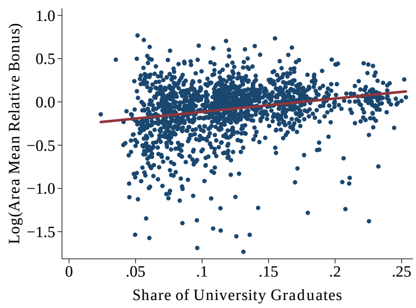
(b) Relative Hourly Wage 2



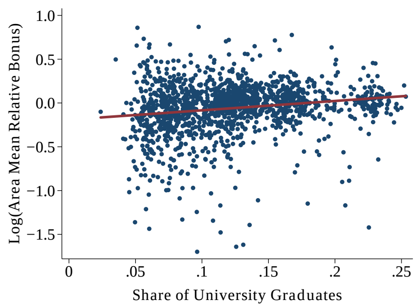
(c) Relative Hourly Wage 3



(d) Relative Bonuses 1



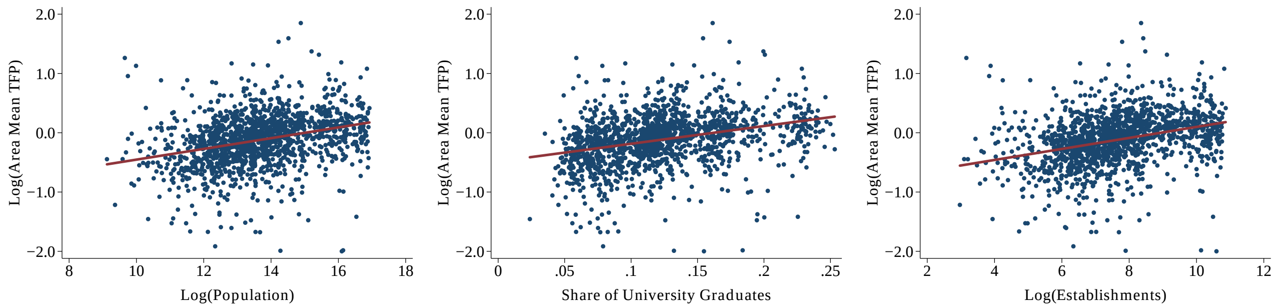
(e) Relative Bonuses 2



(f) Relative Bonuses 3

Figure O.2: Wage and Human Capital Externalities

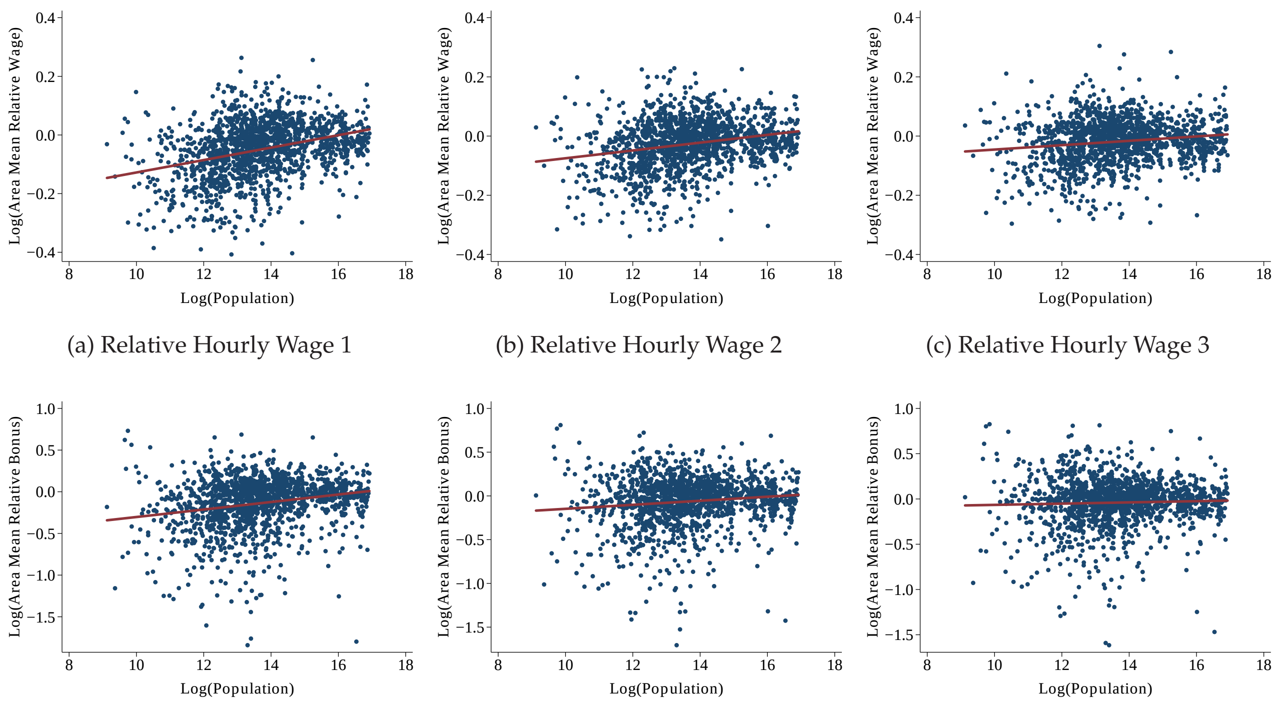
Note: Created by author. Area-year mean wages and share of university graduates are, respectively, averaged across years as $\overline{w}_a^{\text{Type}} = 1/T_a \sum_t \overline{w}_{at}^{\text{Type}}$, and $\overline{\text{Univ}}_a = 1/T_a \sum_t \overline{\text{Univ}}_{at}$, where T_a is the number of years for municipality a observed in the sample.



(a) Population (b) Share of University Graduates (c) Establishments

Figure O.3: TFP and Agglomeration

Note: Created by author. Area-year mean TFP, population, share of university graduates, and the number of manufacturing establishments are respectively, averaged across years as $\overline{\text{TFP}}_a = 1/T_a \sum_t \overline{\text{TFP}}_{at}$, $\overline{\text{Pop}}_a = 1/T_a \sum_t \overline{\text{Pop}}_{at}$, $\overline{\text{Univ}}_a = 1/T_a \sum_t \overline{\text{Univ}}_{at}$, and $\overline{\text{Estb}}_a = 1/T_a \sum_t \overline{\text{Estb}}_{at}$, where T_a is the number of years for municipality a observed in the sample.



(a) Relative Hourly Wage 1

(b) Relative Hourly Wage 2

(c) Relative Hourly Wage 3

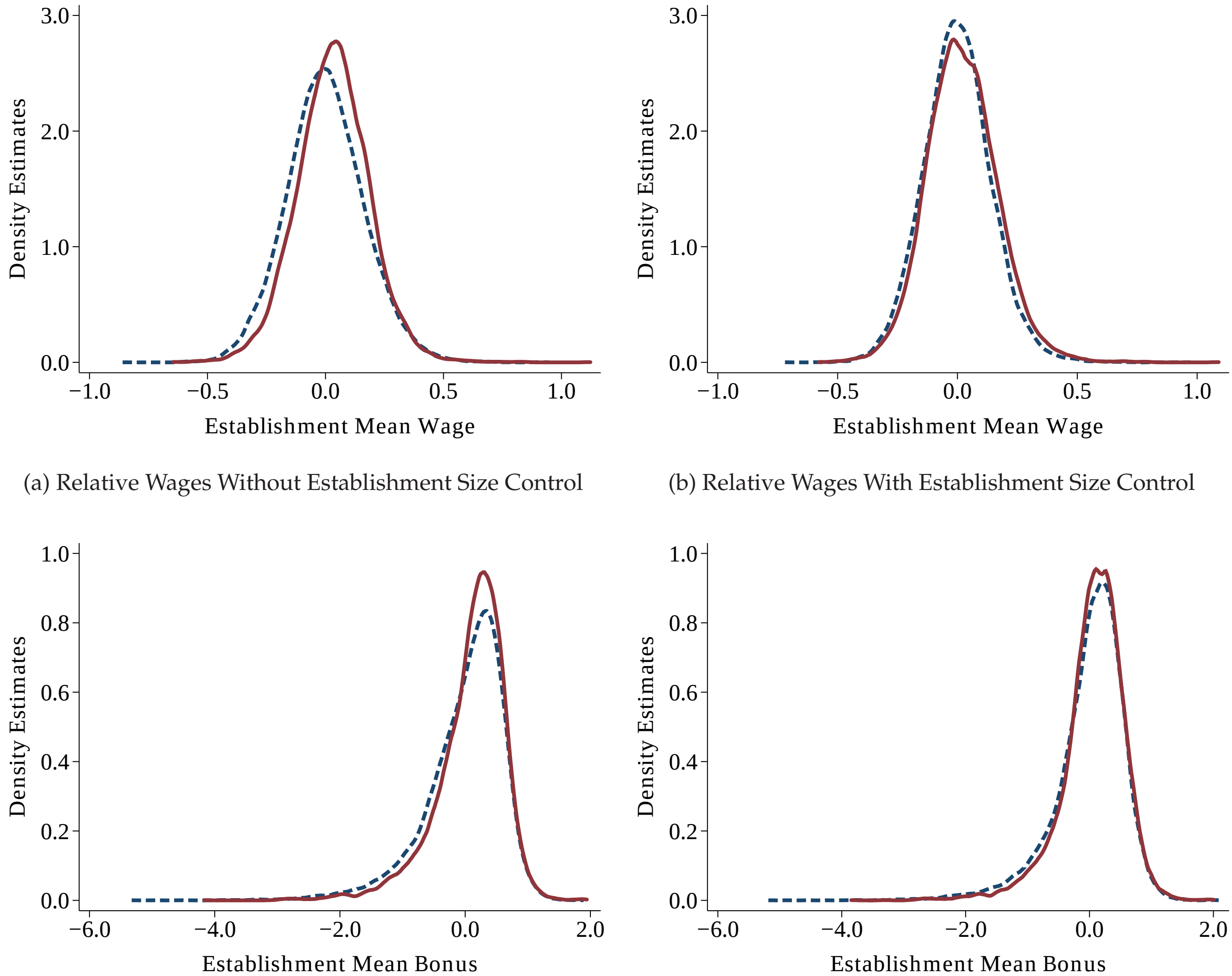
(d) Relative Bonuses 1

(e) Relative Bonuses 2

(f) Relative Bonuses 3

Figure O.4: Wage Controlled for Dynamic Benefits from Agglomeration Economies and City Size

Note: Created by author. Area-year mean wages and manufacturing establishments are, respectively, averaged across years as $\overline{w}_a^{\text{Type}} = 1/T_a \sum_t \overline{w}_{at}^{\text{Type}}$, $\overline{\text{Estb}}_a = 1/T_a \sum_t \overline{\text{Estb}}_{at}$, where T_a is the number of years for municipality a observed in the sample.



(a) Relative Wages Without Establishment Size Control

(b) Relative Wages With Establishment Size Control

(c) Relative Bonuses Without Establishment Size Control

(d) Relative Bonuses With Establishment Size Control

Figure O.5: Spatial Distributions of Establishment Mean Wages and Bonuses Controlled for Dynamic Benefits of Agglomeration Economies

Note: Created by author. Establishment mean wages and bonuses are averaged across years as $\overline{w}_j^{\text{Total1}} = 1/T_j \sum_t \overline{w}_{jt}^{\text{Total1}}$ and $\overline{b}_j^{\text{Total1}} = 1/T_j \sum_t \overline{b}_{jt}^{\text{Total1}}$, where T_j is the number of years for establishment j observed in the sample. Individual characteristics of workers in establishments are controlled for in all panels. Establishment size (employment size and financial capital) is additionally controlled for in Panels (b) and (d), compared to Panels (a) and (c). The red solid (blue dashed) line indicates the establishment mean wage and bonus distributions of cities with above-75 percentile (below-75 percentile) population.