Discussion Paper Series

RIEB

Kobe University

DP2024-09

Distinguishing the Urban Wage Premium from Human Capital Externalities: Evidence from Mexico

Keisuke KONDO

Revised April 11, 2024

Research Institute for Economics and Business Administration

Kobe University

2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Distinguishing the Urban Wage Premium from Human Capital Externalities: Evidence from Mexico

Keisuke Kondo[†]
RIETI & RIEB, Kobe University

Version: April 10, 2024

Abstract

This study bridges the gap between the urban wage premium and human capital externalities in Mexico. High-skilled workers tend to be concentrated in large cities, leading to higher wages in large cities. Merging worker-level microdata with geographical data in Mexico and employing the two-step approach of the Mincer wage equation, this study identifies whether urban wage premium and human capital externalities explain local wage premium, focusing on worker heterogeneity. This study finds that the spatial sorting of workers and firms and human capital externalities entirely explain the urban wage premium in Mexico. An interesting finding is the heterogeneous effect of human capital externalities on the wages of high- and low-skilled workers. Low-skilled workers benefit from human capital externalities, whereas high-skilled workers do not. However, compared with low-skilled workers, high-skilled workers get more than twice the private return to education anywhere they work. This study provides evidence that locations where high-skilled workers are concentrated can pay higher wages to low-skilled workers due to human capital externalities, compensating for their lower private return to education.

JEL classification: J31, R12, R23
Keywords: Urban wage premium, Human capital externalities, Spatial sorting, Private returns to education, Social returns to education

[†] Corresponding Author. Research Institute of Economy, Trade and Industry. 1-3-1 Kasumigaseki, Chiyoda-ku, Tokyo, 100-8901, Japan. (e-mail: kondo-keisuke@rieti.go.jp).
The views expressed in this study are exclusively those of the author and do not represent those of the author's affiliated organizations, the Research Institute of Economy, Trade and Industry (RIETI) and the Research Institute for Economics and Business Administration (RIEB), Kobe University.

1. Introduction

Why do workers earn higher wages in large cities? Recent studies on urban economics have attempted to answer the question of the urban wage premium. Although it is challenging to uncover the channels and sources of the urban wage premium, there have been advanced studies in recent years. As studied by Combes et al. (2008), higher-skilled workers tend to concentrate in large cities, resulting in higher wages. This is known as the spatial sorting of skilled workers. Combes et al. (2010) demonstrated that spatial sorting explains about half of the urban wage premium obtained by the standard method. According to de la Roca and Puga (2017), the dynamic learning process of working in large cities is another source of the urban wage premium. Workers in large cities gradually increase their productivity, resulting in higher wages. Gould (2007) also revealed that urban work experience increases wages, especially white-collar workers continuously earn higher wages after leaving cities.[1]

This study bridges the gap between two types of literature—urban wage premiums and human capital externalities (Glaeser and Maré, 2001; Heuermann et al., 2010). The literature on human capital externalities has found that the concentration of high-skilled workers generates positive externalities to wages and productivity (Rauch, 1993, Acemoglu and Angrist, 2001, Moretti, 2004a, b). For example, Acemoglu and Angrist (2001) estimated private and external rates of return to education in the US, and using the ordinary least squares (OLS) estimation, they found that the external rate of return to education was 0.073 in 1960–1980. Moretti (2004a) also investigated the social return to education using the proportion of college graduates instead of average years of schooling.[2] However, some studies found no evidence of human capital externalities. For example, Acemoglu and Angrist (2001) pointed out that the external return to education is statistically insignificant and its magnitude is considerably small when the exogenous variation in average schooling is employed. Ciccone and Peri (2006) proposed a new approach based on the change in average education and found no evidence of human capital externalities for wages. Heuermann et al. (2010) conducted a literature review by connecting the urban wage premium with human capital externalities because high-skilled workers tend to concentrate in large cities.

---

[1] See the studies by Matano and Naticchioni (2012), Combes et al. (2012), Kondo (2015), Matano and Naticchioni (2016), Silva and Azzoni (2022), Dauth et al. (2022), Belloc et al. (2023), and Butts et al. (2023) for additional studies on the urban wage premium and agglomeration economies. See the studies by Glaeser (1999), Glaeser and Maré (2001), Gould (2007), Glaeser and Resseger (2010), Baum-Snow and Pavan (2012), Kondo (2017a), and de la Roca and Puga (2017) for dynamic learning in larger cities.

[2] The interaction between agglomeration economies and human capital externalities is also an interesting topic. For example, Ciccone and Hall (1996) and Maré and Graham (2013) investigated whether employment density increases productivity. Moretti (2004c) investigated how human capital externalities increases total factor productivity using the U.S. plant-level data.

Merging worker-level microdata with municipal data in Mexico, this study aims to distinguish between urban wage premiums and human capital externalities. Figure 1 depicts a positive relationship between the average years of schooling and population density at the municipal level using the 2005 and 2010 Population and Housing Censuses (*Conteo de Población y Vivienda 2005* and *Censo de Población y Vivienda 2010*) in Mexico, revealing that high-skilled workers tend to work in large cities. However, the high correlation between both variables leads to an econometric issue in identifying the variables that affect wages. Therefore, an exogenous event is generally required to distinguish between urban agglomeration and the concentration of high-skilled workers (Moretti, 2004b).

An advantage of the Mexican data is that there is a significant geographic variation between urban and human capital factors. High-skilled workers tend to move to border states, whereas the traditional metropolitan area is Mexico City (Krugman and Elizondo, 1996). Figure 2 depicts the geographical distributions of the average years of schooling and population density. High-skilled workers are also concentrated in the border states of Mexico, which are less dense than the Mexico City metropolitan area. This variation allows us to identify the important urban wage premium and human capital externalities.

[Figures 1 and 2]

This study contributes to the literature by demonstrating that the spatial sorting of workers and firms and human capital externalities entirely explain the urban wage premium in Mexico. More importantly, it finds heterogeneous human capital externalities between high- and low-skilled workers. Low-skilled workers benefit more from external returns to education in local labor markets than high-skilled workers. However, this study did not find human capital externalities for high-skilled workers, implying that high-skilled workers are concentrated in better locations that attract firms paying higher wages because they can get a higher private rate of return to education anywhere they work. This finding is strongly related to Frías et al. (2024), who show that exports have a significant positive effect on wage premiums in Mexico. It is suggested that highly skilled workers earn higher wages when they work for exporting firms. Referring to Moretti (2012), this study shows that locations where high-skilled workers are concentrated can pay higher wages to low-skilled workers because of human capital externalities, compensating for their lower private return to education.

The remainder of this paper is organized as follows. Section 2 explains the empirical framework. Section 3 presents the worker-level microdata in Mexico. Section 4 discusses the estimation results. Finally, Section 5 concludes.

2. Empirical Framework

2.1. Two-Step Approach for Factor Decomposition

The empirical strategy embodies two steps based on the studies by Combes et al. (2010), de la Roca and Puga (2017), and Kondo (2017b). First, we estimate a Mincer wage equation with municipality dummy variables η_a to identify a municipal wage premium. Thus, the regression model takes the following form:

$$\log(w_{iat}) = \boldsymbol{X}_{it}^{\text{Individual}}\boldsymbol{\beta} + \boldsymbol{X}_{it}^{\text{Firm}}\boldsymbol{\gamma} + \eta_a + \tau_t + e_{iat}, \quad (1)$$

where w_{iat} is the hourly wage of worker i residing in municipality a at time t; $\boldsymbol{X}_{it}^{\text{Individual}}$ is a vector of individual control variables (years of schooling, years of experience, gender dummy, marriage dummy, occupation dummies, and sector dummies) of worker i at time t; $\boldsymbol{X}_{it}^{\text{Firm}}$ is a vector of firm control variables of worker i at time t (firm size dummies); $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are vectors of the corresponding parameters; τ_t is a quarter dummy; and e_{iat} is an error term. The estimates of municipal dummies $\hat{\eta}_a$ are interpreted as municipal wage premiums. This approach does not control for unobservable individual skills but observable ones, such as education and experience.[3]

This study calculates three types of municipal wage fixed effects based on Equation (1). The first calculates the simple average municipal wage among workers. The regression includes no controls as follows:

$$\begin{aligned} \log(w_{iat}) &= \eta_a + \tau_t + e_{iat}, \\ \log(\overline{w}_a^N) &= \hat{\eta}_a - \bar{\hat{\eta}}_a, \end{aligned} \quad (2)$$

where $\log(w_{iat})$ still includes spatial sorting factors. For example, this average municipal wage becomes higher in large areas if more skilled workers are concentrated in large areas.

The second controls for spatial sorting of workers' skills. The regression includes individual controls ($\boldsymbol{X}_{it}^{\text{Individual}}$) as follows:

[3] Combes et al. (2008), Combes et al. (2011), and de la Roca and Puga (2017) controlled for unobservable individual fixed effects in their estimations. However, municipal dummies can only be identified if workers migrated across municipalities during the research period. This study does not control for unobservable individual fixed effects due to data limitation. The relationship between municipal wage fixed effects and workers' location choices might generate endogeneity bias in the second step, known as spatial sorting (Combes et al., 2008).

$$\log(w_{iat}) = \boldsymbol{X}_{it}^{\text{Individual}}\boldsymbol{\beta} + \eta_a + \tau_t + e_{iat},$$
$$\log(\overline{w}_a^{\text{I}}) = \hat{\eta}_a - \bar{\hat{\eta}}_a, \tag{3}$$

where $\log(\overline{w}_a^{\text{I}})$ still includes firm factors. For example, the average municipal wage is higher in large cities if large firms are concentrated in large cities.

The third controls for spatial sorting of workers' skills and firms because large firms pay higher wages (Brown and Medoff, 1989). The regression includes individual and firm controls ($\boldsymbol{X}_{it}^{\text{Individual}}$ and $\boldsymbol{X}_{it}^{\text{Firm}}$, respectively) as follows:

$$\log(w_{iat}) = \boldsymbol{X}_{it}^{\text{Individual}}\boldsymbol{\beta} + \boldsymbol{X}_{it}^{\text{Firm}}\boldsymbol{\gamma} + \eta_a + \tau_t + e_{iat},$$
$$\log(\overline{w}_a^{\text{IF}}) = \hat{\eta}_a - \bar{\hat{\eta}}_a, \tag{4}$$

where $\log(\overline{w}_a^{\text{IF}})$ strongly captures local factors of municipal wage premium in municipality a.

This study considers the heterogeneous aspects of municipal wage fixed effects. It divides workers into two groups (high- and low-skilled workers) and estimates Equations (**2**)–(**4**) to obtain municipal wage fixed effects η_a.

2.2. Urban Wage Premium and Human Capital Externalities

To distinguish the urban wage premium from human capital externalities, we regress the three types of estimated municipal fixed effects on population density and average years of schooling in the second step. Based on the study by Glaeser and Resseger (2010), the regression takes the following form:

$$\log(w_{iat}) = \text{Const} + \lambda \log(\text{Dens}_a) + \phi \text{Educ}_a + \varepsilon_a, \tag{5}$$

where $\log(\overline{w}_a^{\text{Type}}) \in \{\log(\overline{w}_a^{\text{N}}), \log(\overline{w}_a^{\text{F}}), \log(\overline{w}_a^{\text{IF}})\}$ is the wage premium of municipality a; Const is a constant term; Dens_a is the population density of municipality a; Educ_a is the average years of schooling of municipality a; and ε_a is an error term.[4]

Our interests are in two parameters λ and ϕ, which measure the urban wage premium and external rate of return to education, respectively. First, this study estimates Equation (5) in terms of urban wage premium and human capital externalities separately. Next, it considers the simultaneous inclusion of both variables.

An econometric issue is an endogeneity concern. Following Rauch (1993), this study considers an omitted

[4] This study additionaly considers the population potential (Stewart, 1947). See Section B in Supplementary Information.

variable bias. The location choice of firms also affects the wage level. Rauch (1993) found higher rents in large cities with higher education levels. The areas with better residence and business conditions attract more productive and profitable firms that pay higher wages. As a result, high-skilled workers concentrate in better locations, generating a positive relationship between local wages and average years of schooling. As a robustness check, this study introduces the number of housing units with home electrical appliances to control residence and business conditions.

3. Data

This study uses the microdata on Mexican workers from the National Occupation and Employment Survey (*Encuesta Nacional de Ocupación y Empleo*, ENOE). The survey has been conducted quarterly by the National Institute of Statistics and Geography (*Instituto Nacional de Estadística y Geografía*, INEGI) since 2005; ENOE compiles workers' characteristics (e.g., age, gender, place of residence, and marriage), hourly wage, firm size, occupation, and sector. This study merges ENOE microdata from 2005:Q1 to 2010:Q4 with the municipal data of the 2005 and 2010 Population and Housing Censuses.

Table 1 presents the descriptive statistics for the first-step regression. The dependent variable, hourly wage, is adjusted using the consumer price index published by the INEGI. The base period is 2005:Q1. Spatial differences in cost of living are adjusted using area differences in minimum wages. As of 2005, there were three categories of minimum wage in Mexico depending on the degree of urbanization. The base area is the lowest minimum wage of the three categories. In line with the literature on the Mincer wage equation, we calculate years of experience by subtracting six years and years of schooling from age.[5] The regression models include control variables on gender, marriage status, occupation, and firm size. We create occupation dummies at the two-digit level based on the occupation code. Firm-size dummies are created for 1–10, 11–20, 21–50, 51–100, 101–250, 251–500, and 501 workers and over (the reference is Category 1–10 workers).

[Table 1]

The second-step regression includes municipal variables. This study considers the population density and average years of schooling at the municipal level. There are 2,454 municipalities in the 2005 census. Population density is calculated from population and area data obtained from the 2005 and 2010 censuses.[6] The average years of schooling for people aged 15 and above are also available at the municipal level in the

[5] When years of experience take on a negative value, we drop the observations from the sample.

[6] The National System of Municipal Information (*Sistema Nacional de Información municipal*, SNIM) provided summarized municipal data, indicating the area, latitude, and longitude. Currently, the municipal data are available on the INEGI website, https://www.inegi.org.mx/programas/ccpv/2005/ (accessed on March 22, 2024)

2005 and 2010 Population and Housing Censuses.

This study uses spatial smoothing based on the study by de la Roca and Puga (2017) to avoid bias arising from administrative border discontinuity. Population density is calculated as population divided by area (in km^2) within a radius of 10 km from the administrative center of each municipality (*cabecera municipal*). The average years of schooling are also smoothed by the sum of average years of schooling divided by the number of municipalities within a radius of 10 km from the administrative center of each municipality. Based on the latitude and longitude, this study calculates the great circle distances between any two municipalities using Vincety's formula (Vincenty, 1975). Finally, municipal wage fixed effects obtained from the first-step estimation of the Mincer wage equation are matched with population density and average years of schooling via municipal code.

Table 2 presents the descriptive statistics for the second-step regression. The number of municipalities becomes 653 for total workers in the second-step regression. As a robustness check, this study includes the number of housing units with home electrical appliances in each municipality, capturing the local residence and business conditions. This variable is also available from the 2005 and 2010 Population and Housing Censuses.

[Table 2]

4. Estimation Results and Discussions

4.1. First-Step Estimation Results

Table 3 presents the first-step estimation results of the Mincer wage equation with municipality dummies by worker skill. Total, high-skilled, and low-skilled workers correspond to Columns (1)–(3), Columns (4)–(5), and Columns (6)–(9) of Table 3, respectively. The wages in Equations (2)–(4) correspond to each column of Table 3. The estimated municipal fixed effects are used in the second-step regressions.

To determine whether our estimation results are reliable, we compare them in Column (3) of Table 3 with those obtained by Chiquiar (2008). He estimated a Mincer wage equation using data from the 1990 and 2000 censuses, obtaining 0.051 for years of education; 0.018 for years of experience; −0.023 for years of experience squared divided by 100; and 0.085 for a marriage dummy using the 2000 census data after controlling for observable workers' characteristics, region, position, occupation, sector, site-specific feature, and globalization. The results are similar to the estimates obtained in Table 3. Although we use microdata from ENOE, parameters in our regression model seem to be correctly estimated.

This study finds heterogeneous private rates of return to education between high- and low-skilled workers.

While it is 8.1% for high-skilled workers, it is 3.2% for low-skilled workers, implying that a college degree is a key factor for wage determination in Mexico. A firm-size dummy also indicates the heterogeneous effects on wages of high- and low-skilled workers. Large firms pay higher salaries for both workers, and high-skilled workers earn higher wages in large firms than low-skilled firms.

[Table 3]

4.2. Second-Step Estimation Results

Figure 3 compares the distributions of municipal wage fixed effects between municipalities with above-median (below-median) population density by worker skill. In Panels (a) to (c) of Figure 3, both distributions overlap closely by controlling for spatial sorting of worker and firm factors. Visually, we hardly find evidence of urban wage premium after controlling for spatial sorting, which is statistically tested later.

[Figure 3]

Figure 4 compares the distributions of municipal wage fixed effects between municipalities with above- and below-median average years of schooling by worker skill. In Panels (a) to (c) of Figure 4, the gap between both distributions becomes smaller by controlling for spatial sorting of worker and firm factors. Both distributions closely overlap for high-skilled workers in Panel (f) of Figure 4, whereas both distributions do not overlap for low-skilled workers, implying that the concentration of high-skilled workers increases individual wages of low-skilled workers at the regional level.

[Figure 4]

Figure 5 illustrates urban wage premium and the external return to education after controlling for spatial sorting. Panels (a) and (b) of Figure 5 correspond to the urban wage premiums for high- and low-skilled workers. Panels (c) and (d) of Figure 5 correspond to human capital externalities for high- and low-skilled workers. Compared with the urban wage premium, this study finds a clear positive relationship between the estimated municipal wage fixed effect and average years of schooling, implying that human capital externalities are important, especially for low-skilled workers.

[Figure 5]

Table 4 presents the estimation results of the second-step regression by worker skill. Columns (1)–(3), (4)–(6), and (7)–(9) of Table 4 correspond to total, high-skilled, and low-skilled workers, respectively. Table 3 includes two regression models, where population density and average years of schooling are included separately.

[Table 4]

The urban wage premium is confirmed in the upper part of Table 4. When the spatial sorting is not controlled, the elasticity of population density to wage premium is 0.063 in Column (1) of Table 4. When the spatial sorting of workers and firms is controlled, the elasticity of population density to wage premium is 0.031 in Column (3) of Table 4, which is less than half of that of Column (1) of Table 4. Furthermore, the elasticity of population density to wage premium is 0.018 for high-skilled workers and 0.033 for low-skilled workers, implying that the urban wage premium is heterogeneous across workers.

The magnitude of the urban wage premium is close to that of previous studies. For example, Combes et al. (2008) obtained an elasticity of 0.030 in France, and similarly, Combes et al. (2010) obtained an elasticity of 0.027 in France. Furthermore, de la Roca and Puga (2017) obtained an elasticity of 0.025 in Spain; D'Costa and Overman (2014) obtained an elasticity of 0.022 in the UK; and Kondo (2017b) obtained an elasticity of 0.13 in Japan after controlling for firm factors. Using Italian microdata, Mion and Naticchioni (2009) obtained an elasticity of 0.007 after controlling for individual fixed effects. Applying a quantile regression approach to Italian microdata, Matano and Naticchioni (2012) obtained a range of 0.01–0.02 for the industrial and service sectors.

The lower part of Table 4 confirms the human capital externalities. When spatial sorting is not controlled, the elasticity of average years of schooling to wage premium is 0.103 in Column (1) of Table 4. When the spatial sorting of workers and firms is controlled, the elasticity of average years of schooling to wage premium is 0.047 in Column (3) of Table 4, which is less than half in Column (1) of Table 4. The magnitude is close to that of Acemoglu and Angrist (2001), indicating that the external rate of return to education ranges from 0.061 to 0.073 using the OLS estimation. After controlling for spatial sorting, this study finds heterogeneous human capital externalities as found earlier. The external return to education is 0.030 for high-skilled workers (Column (6) of Table 3) and 0.051 for low-skilled workers (Column (9) of Table 3).

Table 5 presents the estimation results of the second-step regression by worker skill by distinguishing the urban wage premium from human capital externalities. Columns (1)–(3), (4)–(6), and (7)–(9) of Table 5 correspond to total, high-skilled, and low-skilled workers, respectively. The estimation results indicate that the population density is not statistically significant when the average years of schooling are included in the regression, implying that the concentration of high-skilled workers further leads to positive externalities in local labor markets. An entire part of the urban wage premium is explained by the external return to education. As mentioned earlier, this study finds heterogeneous human capital externalities between high- and low-skilled workers. The magnitude of human capital externalities slightly decreases in each column compared with those in Table 4.

[Table 5]

The summary of the OLS estimation is as follows. This study uncovers two sources of urban wage premiums. Spatial sorting is the first source of the urban wage premium, which leads to higher wages in large cities. Furthermore, the concentration of high-skilled workers generates positive externalities in local labor markets, which is the second source of the urban wage premium. This study uncovers that spatial sorting of workers and firms and human capital externalities entirely explain the urban wage premium in Mexico. It emphasizes the heterogeneous effects of human capital externalities on wages. Low-skilled workers enjoy more benefits of external return to education, whereas the external rate of return to education of high-skilled workers is relatively small and less than half of that of low-skilled workers. Considering endogeneity issues, this study further examines the magnitude of human capital externalities for high- and low-skilled workers.

4.3. Robustness Check

Table 6 presents the robustness check of the estimation results. Following Rauch (1993), this study considers that the omitted variable generates a bias for urban wage premium and human capital externalities. Table 6 includes an additional control variable—the number of housing units with home electrical appliances—denoting local residence and business conditions (See also Section A in Supplementary Information).

This study finds that the variable on local residence and business conditions has a statistically significant effect on average municipal wages. The magnitude does not change after controlling for spatial sorting of individual worker skills, implying that this variable captures localized factors. For example, better locations attract productive firms paying higher wages, and, as a result, high-skilled workers tend to concentrate there. The robustness check confirms this mechanism, implying that human capital externalities are overestimated.

This study finds heterogeneous human capital externalities between high- and low-skilled workers. The external return to education is statistically insignificant and 0.8% for high-skilled workers in Column (6) of Table 6, implying that high-skilled workers do not benefit from the external return to education. This is consistent with previous studies, such as those by Acemoglu and Angrist (2001) and Ciccone and Peri (2006). Our estimation results reveal that high-skilled workers are concentrated in better locations that attract firms paying higher wages because their private return to education is sufficiently high. Thus, they can increase their wages through individual human capital accumulation.

This study finds that only low-skilled workers benefit from the external return to education. For them, the external rate of return to education is statistically significant and is 2.7% after controlling for spatial sorting and firm location factors. This rate is sufficiently high because their private return to education is 3.1% in

Column (9) of Table 3. This finding implies that locations where high-skilled workers are concentrated attract low-skilled workers because the local average wage rate rises at a sufficiently high rate. As discussed by Moretti (2012), this study provides evidence of static human capital externalities for low-skilled workers.

[Table 6]

5. Conclusion

This study has investigated sources of the urban wage premium, focusing on human capital externalities. Merging worker-level microdata with the geographical data of the censuses in Mexico and employing the two-step approach of the Mincer wage equation, it distinguishes between urban wage premium and human capital externalities because high-skilled workers tend to concentrate in large cities. A novel feature of this study is to uncover how a worker's skill level affects the magnitude of the urban wage premium and human capital externalities.

This study found that the spatial sorting of workers and firms and human capital externalities entirely explain the urban wage premium in Mexico. An interesting finding is the heterogeneous effect of human capital externalities between high- and low-skilled workers, which is consistent with the findings of Moretti (2004a). Only low-skilled workers benefit from external returns to education in local labor markets. The external rate of return to education for low-skilled workers is estimated as 5.0% with the OLS estimation and 2.7% after controlling for local residence and business conditions. For high-skilled workers, the external rate of return to education is estimated as 3.1% with the OLS estimation and 0.3% after controlling for local residence and business conditions. This study found that, despite where they work, the private rate of return to education for high-skilled workers is more than twice that of low-skilled workers, implying that high-skilled workers do not care about the external return to education because they increase their wages through individual human capital investment. These findings are also consistent with the studies by Acemoglu and Angrist (2001) and Ciccone and Peri (2006), who found no evidence of human capital externalities.

This study has several limitations. First, it did not control for unobservable individual fixed effects owing to data limitation. Spatial sorting requires a rich panel dataset with migration information. Second, this study did not focus on the dynamic process of learning by working in large cities. The estimation results imply that high-skilled workers search for locations for individual human capital investment, whereas low-skilled workers search for locations with higher external returns to education. Third, this study did not directly control for spatial sorting of firms owing to data limitation. Employer–employee matched data are required to consider how firm characteristics and location factors affect wages (Abowd and Kramarz, 1999, Abowd et al., 1999, Dauth et al., 2022, Frías et al, 2024). Although this study did not find external return to education

for high-skilled workers at the regional level, it is also possible to examine external return to education at the firm level. Although the estimation results are valid, resolving these issues will uncover the detailed mechanisms of the urban wage premium and human capital externalities.

References

Abowd JM, Kramarz F (1999) The analysis of labor markets using matched employer–employee data. In: Ashenfelter O, Card D (eds) Handbook of Labor Economics, vol 3, Part B. Elsevier, Amsterdam, chap 40, p 2629–2710, https://doi.org/10.1016/S1573-4463(99)30026-2

Abowd JM, Kramarz F, Margolis DN (1999) High wage workers and high wage firms. Econometrica 67(2):251–333. https://doi.org/10.1111/1468-0262.00020

Acemoglu D, Angrist J (2001) How large are human-capital externalities? Evidence from compulsory-schooling laws. In: Bernanke BS, Rogoff K (eds) NBER Macroeconomics Annual 2000, vol 15. MIT Press, chap 1, p 9–74, https://doi.org/10.1086/654403

Baum-Snow N, Pavan R (2012) Understanding the city size wage gap. Review of Economic Studies 79(1):88–127. https://doi.org/10.1093/restud/rdr022

Belloc M, Naticchioni P, Vittori C (2023) Urban wage premia, cost of living, and collective bargaining. Journal of Economic Geography 23(1):25–50. https://doi.org/10.1093/jeg/lbab048

Brown C, Medoff J (1989) The employer size-wage effect. Journal of Political Economy 97(5):1027– 1059. https://doi.org/10.1086/261642

Butts K, Jaworski T, Kitchens C (2023) The urban wage premium in historical perspective. NBER Working Paper No. 31387, https://doi.org/10.3386/w31387

Chiquiar D (2008) Globalization, regional wage differentials and the Stolper-Samuelson theorem: Evidence from Mexico. Journal of International Economics 74(1):70–93. https://doi.org/10.1016/j.jinteco.2007.05.009

Ciccone A, Hall RE (1996) Productivity and the density of economic activity. American Economic Review 86(1):54–70. https://doi.org/https://www.jstor.org/stable/2118255

Ciccone A, Peri G (2006) Identifying human-capital externalities: Theory with applications. Review of Economic Studies 73(2):381–412. https://doi.org/10.1111/j.1467-937X.2006.00380.x

Combes PP, Duranton G, Gobillon L (2008) Spatial wage disparities: Sorting matters! Journal of Urban Economics 63(2):723–742. https://doi.org/10.1016/j.jue.2007.04.004

Combes PP, Duranton G, Gobillon L, et al (2010) Estimating agglomeration economies with history, geology, and worker effects. In: Glaeser EL (ed) Agglomeration Economics. University of Chicago Press, Chicago,

chap 1, p 15–66

Combes PP, Duranton G, Gobillon L (2011) The identification of agglomeration economies. Journal of Economic Geography 11(2):253–266. https://doi.org/10.1093/jeg/lbq038

Combes PP, Duranton G, Gobillon L, et al (2012) Sorting and local wage and skill distributions in France. Regional Science and Urban Economics 42(6):913–930. https://doi.org/10.1016/j.regsciurbeco.2012.11.003

Dauth W, Findeisen S, Moretti E, et al (2022) Matching in cities. Journal of the European Economic Association 20(4):1478–1521. https://doi.org/10.1093/jeea/jvac004

D'Costa S, Overman HG (2014) The urban wage growth premium: Sorting or learning? Regional Science and Urban Economics 48(C):168–179. https://doi.org/10.1016/j.regsciurbeco.2014.06.006

de la Roca J, Puga D (2017) Learning by working in big cities. Review of Economic Studies 84(1):106–142. https://doi.org/10.1093/restud/rdw031

Frías JA, Kaplan DS, Verhoogen E, Alfaro-Serrano D (2024) Exports and wage premiums: Evidence from Mexican employer-employee data. The Review of Economics and Statistics 106(2):305–321. https://doi.org/10.1162/rest_a_01178

Glaeser EL (1999) Learning in cities. Journal of Urban Economics 46(2):254–277. https://doi.org/10.1006/juec.1998.2121

Glaeser EL, Maré DC (2001) Cities and skills. Journal of Labor Economics 19(2):316–342. https://doi.org/10.1086/319563

Glaeser EL, Resseger MG (2010) The complementarity between cities and skills. Journal of Regional Science 50(1):221–244. https://doi.org/10.1111/j.1467-9787.2009.00635.x

Gould ED (2007) Cities, workers, and wages: A structural analysis of the urban wage premium. Review of Economic Studies 74(2):477–506. https://doi.org/10.1111/j.1467-937X.2007.00428.x

Heuermann D, Halfdanarson B, Süedekum J (2010) Human capital externalities and the urban wage premium: Two literatures and their interrelations. Urban Studies 47(4):749–767. https://doi.org/10.1177/0042098009352363

Kondo K (2015) Disentangling the channels of agglomeration impacts on individual wages: The case of Mexican banking sector. Mimeo

Kondo K (2017a) Dynamic benefits of working in large cities: Evidence from Japanese matched employer–employee data. RIETI Discussion Paper No. 17-E-043

Kondo K (2017b) Urban wage premium revisited: Evidence from Japanese matched employer–employee data. RIETI Discussion Paper No. 17-E-047

Krugman P, Elizondo RL (1996) Trade policy and the third world metropolis. Journal of Development Economics 49(1):137–150. https://doi.org/10.1016/0304-3878(95)00055-0

Maré D, Graham D (2013) Agglomeration elasticities and firm heterogeneity. Journal of Urban Economics 75(C):44–56. https://doi.org/10.1016/j.jue.2012.12.002

Matano A, Naticchioni P (2012) Wage distribution and the spatial sorting of workers. Journal of Economic Geography 12(2):379–408. https://doi.org/10.1093/jeg/lbr013

Matano A, Naticchioni P (2016) What drives the urban wage premium? Evidence along the wage distribution. Journal of Regional Science 56(2):191–209. https://doi.org/10.1111/jors.12235

Mion G, Naticchioni P (2009) The spatial sorting and matching of skills and firms. Canadian Journal of Economics 42(1):28–55. https://doi.org/10.1111/j.1540-5982.2008.01498.x Moretti E (2004a) Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. Journal of Econometrics 121(1–2):175–212. https://doi.org/10.1016/j.jeconom.2003.10.015

Moretti E (2004b) Human capital externalities in cities. In: Henderson JV, Thisse JF (eds) Handbook of Regional and Urban Economics, vol 4. Elsevier, Amsterdam, chap 51, p 2243–2291, https://doi.org/10.1016/S1574-0080(04)80008-7

Moretti E (2004c) Workers' education, spillovers, and productivity: Evidence from plant-level production functions. American Economic Review 94(3):656–690. https://doi.org/10.1257/0002828041464623

Moretti E (2012) The new geography of jobs. Houghton Mifflin Harcourt, Boston, MA

Rauch JE (1993) Productivity gains from geographic concentration of human capital: Evidence from the cities. Journal of Urban Economics 34(3):380–400. https://doi.org/10.1006/juec.1993.1042

Silva D, Azzoni C (2022) Worker and firm heterogeneity, agglomeration, and wages in Brazil. Papers in Regional Science 101(1):107–134. https://doi.org/10.1111/pirs.12637

Stewart JQ (1947) Empirical mathematical rules concerning the distribution and equilibrium of population. Geographical Review 37(3):461–475. https://doi.org/10.2307/211132

Vincenty T (1975) Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. Survey Review 23(176):88–93. https://doi.org/10.1179/sre.1975.23.176.88

Author Contributions

Keisuke Kondo: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Funding acquisition.

Declaration of Competing Interest

The author declares no competing interests.

Funding

This study was supported by JSPS KAKENHI Grant Numbers 21K01497.

Acknowledgments

I thank Kensuke Teshima for his helpful comments and suggestions. All the remaining errors are naturally my own. The original study was conducted under a scholarship granted by the Government of Mexico through the Ministry of Foreign Affairs of Mexico in 2012–2013. I am grateful for the benefits received during my stay at the Instituto Tecnológico Autónomo de México in 2012–2013.

Supplementary Information

The supplementary information is available.

Table 1 Descriptive Statistics of the First-Step Regression

Variable	Mean	S.D.	Min	Median	Max
Hourly Wage (Pesos)	25.957	20.835	1.695	19.380	213.333
Years of Schooling	9.577	4.247	0.000	9.000	17.000
Years of Experience	19.154	13.724	0.000	17.250	65.750
Years of Experience in Firm	6.817	7.638	0.000	3.750	45.000
Dummy (1=Female)	0.384	0.486	0.000	0.000	1.000
Dummy (1=Married)	0.451	0.498	0.000	0.000	1.000
Firm-Size Dummy (1–10 workers)	0.478	0.500	0.000	0.000	1.000
Firm-Size Dummy (11–50 workers)	0.236	0.425	0.000	0.000	1.000
Firm-Size Dummy (51–250 workers)	0.070	0.255	0.000	0.000	1.000
Firm-Size Dummy (251–500 workers)	0.046	0.210	0.000	0.000	1.000
Firm-Size Dummy (501 workers or more)	0.170	0.375	0.000	0.000	1.000
Occupation 1 Dummy	0.084	0.277	0.000	0.000	1.000
Occupation 2 Dummy	0.048	0.213	0.000	0.000	1.000
Occupation 3 Dummy	0.016	0.124	0.000	0.000	1.000
Occupation 4 Dummy	0.137	0.344	0.000	0.000	1.000
Occupation 5 Dummy	0.279	0.448	0.000	0.000	1.000
Occupation 6 Dummy	0.115	0.319	0.000	0.000	1.000
Occupation 7 Dummy	0.058	0.233	0.000	0.000	1.000
Occupation 8 Dummy	0.168	0.374	0.000	0.000	1.000
Occupation 9 Dummy	0.032	0.175	0.000	0.000	1.000
Occupation 10 Dummy	0.064	0.245	0.000	0.000	1.000
Occupation 11 Dummy	0.000	0.003	0.000	0.000	1.000
Sector 1 Dummy	0.069	0.254	0.000	0.000	1.000
Sector 2 Dummy	0.015	0.121	0.000	0.000	1.000
Sector 3 Dummy	0.169	0.375	0.000	0.000	1.000
Sector 4 Dummy	0.092	0.290	0.000	0.000	1.000
Sector 5 Dummy	0.153	0.360	0.000	0.000	1.000
Sector 6 Dummy	0.066	0.247	0.000	0.000	1.000
Sector 7 Dummy	0.056	0.230	0.000	0.000	1.000
Sector 8 Dummy	0.057	0.233	0.000	0.000	1.000
Sector 9 Dummy	0.105	0.307	0.000	0.000	1.000
Sector 10 Dummy	0.128	0.334	0.000	0.000	1.000
Sector 11 Dummy	0.089	0.285	0.000	0.000	1.000

Notes: The number of observations is 1,973,517 from the 2005Q1–2010Q4 ENOE survey. Hourly wage is de ﬂ ated by the quarterly average monthly consumer price index (2010=100) and the minimum wages of the municipalities where workers live. Occupation 1: Professionals, Engineers, and Artists; Occupation 2: Education workers; Occupation 3: Financiers and Directors; Occupation 4: O ﬃ ce workers; Occupation 5: Blue collar workers, Craft workers, and helpers; Occupation 6: Sellers; Occupation 7: Transport operators; Occupation 8: Personal service workers; Occupation 9: Security o ﬃ cers; Occupation 10: Farmers; Occupation 11: Not speci ﬁ ed. Sector 1: Agriculture, livestock, silviculture, hunting, and ﬁ shing; Sector 2: Mining and electricity; Sector 3: Manufacture; Sector 4: Manufacture; Sector 5: Manufacture; Sector 6: Restaurant and lodging; Sector 7: Transport, communication, postal services, and warehouse; Sector 8: Professional, ﬁ nancial, and corporate services; Sector 9: Social services; Sector 10: Various services; Sector 11: Government and International Organizations.

Table 2 Descriptive Statistics for Second-Step Regression

Variables	Obs.	Mean	S.D.	Min	Median	Max
Total Workers						
Municipal Wage FE (No Control)	527	−0.192	0.257	−1.349	−0.182	1.033
Municipal Wage FE (Individual Controls)	527	−0.117	0.186	−1.025	−0.111	0.562
Municipal Wage FE (Individual & Firm Controls)	527	−0.106	0.178	−0.994	−0.096	0.507
High-Skilled Workers						
Municipal Wage FE (No Control)	481	−0.088	0.320	−1.617	−0.054	1.095
Municipal Wage FE (Individual Controls)	481	−0.105	0.226	−1.021	−0.098	1.226
Municipal Wage FE (Individual & Firm Controls)	481	−0.081	0.217	−1.041	−0.075	1.156
Low-Skilled Workers						
Municipal Wage FE (No Control)	527	−0.145	0.230	−1.227	−0.129	0.867
Municipal Wage FE (Individual Controls)	527	−0.119	0.195	−1.038	−0.105	0.671
Municipal Wage FE (Individual & Firm Controls)	527	−0.109	0.188	−1.010	−0.095	0.616
Area Variables						
Average Years of Schooling within 10km	527	7.254	1.396	3.400	7.158	10.980
log(Population Density within 10km)	527	4.044	1.792	−1.452	3.996	9.087
log(Population Potential; delta = 2)	527	7.888	1.481	4.680	7.781	13.356
Log(Housing with H.E.A. within 10km)	527	7.579	1.747	2.635	7.499	12.101

Notes: The unit of observation is the municipality. H.E.A. stands for home electrical appliances. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and working experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

Table 3 First-Step Estimation Results of the Mincer Wage Equation

	Dependent Variable: Hourly Wage, $\log(w_{iat})$								
	Total Workers			High-Skilled Workers			Low-Skilled Workers		
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years of Schooling		0.0477***	0.0452***		0.0800***	0.0757***		0.0338***	0.0314***
		(0.0009)	(0.0009)		(0.0027)	(0.0026)		(0.0008)	(0.0007)
Years of Experience		0.0184***	0.0175***		0.0263***	0.0248***		0.0181***	0.0173***
		(0.0003)	(0.0002)		(0.0008)	(0.0008)		(0.0004)	(0.0004)
Years of Experience Squared (× 1/100)		−0.0226***	−0.0213***		−0.0439***	−0.0411***		−0.0241***	−0.0230***
		(0.0004)	(0.0004)		(0.0020)	(0.0018)		(0.0006)	(0.0006)
Dummy (1=Female)		−0.1000***	−0.0969***		−0.0614***	−0.0524***		−0.1063***	−0.1041***
		(0.0030)	(0.0033)		(0.0046)	(0.0043)		(0.0044)	(0.0048)
Dummy (1=Married)		0.0754***	0.0714***		0.0924***	0.0888***		0.0706***	0.0667***
		(0.0019)	(0.0019)		(0.0038)	(0.0037)		(0.0024)	(0.0024)
Firm-Size Dummy (11–50 workers)			0.0909***			0.1317***			0.0873***
			(0.0031)			(0.0062)			(0.0052)
Firm-Size Dummy (51–250 workers)			0.1158***			0.2017***			0.0992***
			(0.0059)			(0.0074)			(0.0105)
Firm-Size Dummy (251–500 workers)			0.1383***			0.2349***			0.1186***
			(0.0072)			(0.0093)			(0.0130)
Firm-Size Dummy (501 workers or more)			0.1975***			0.2658***			0.1838***
			(0.0111)			(0.0123)			(0.0189)
Occupation Dummy	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Sector Dummy	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Quarterly Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed-Effect ($\hat{\eta}_a$)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	1,973,517	1,973,517	1,973,517	349,304	349,304	349,304	1,624,213	1,624,213	1,624,213
Number of Municipalities	1,207	1,207	1,207	1,068	1,068	1,068	1,207	1,207	1,207
Adjusted R^2	0.1006	0.4400	0.4469	0.0538	0.3335	0.3468	0.1143	0.3455	0.3528

Notes: Heteroskedasticity-consistent standard errors clustered at the municipal level are in parenthesis. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and working experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

Table 4 Second-Step Estimation Results of Urban Wage Premium and Human Capital Externalities by Workers' Skill

	Dependent Variable: Municipal Wage Fixed-Effect, $\log\left(\bar{w}_a^{\text{Type}}\right)$								
	Total Workers			High-Skilled Workers			Low-Skilled Workers		
	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Regression Model 1: Urban Wage Premium								
Log(Population Density)	0.0632***	0.0347***	0.0310***	0.0135**	0.0267***	0.0183***	0.0480***	0.0360***	0.0327***
	(0.0109)	(0.0062)	(0.0061)	(0.0057)	(0.0050)	(0.0049)	(0.0080)	(0.0064)	(0.0063)
State Fixed-Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Municipalities	527	527	527	481	481	481	527	527	527
Adjusted R^2	0.4517	0.6686	0.6604	0.1990	0.4279	0.3899	0.5613	0.6921	0.6865
	Regression Model 2: Human Capital Externalities								
Average Years of Schooling	0.1034***	0.0531***	0.0474***	0.0237**	0.0425***	0.0299***	0.0778***	0.0555***	0.0506***
	(0.0105)	(0.0068)	(0.0067)	(0.0100)	(0.0078)	(0.0074)	(0.0083)	(0.0064)	(0.0064)
State Fixed-Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Municipalities	527	527	527	481	481	481	527	527	527
Adjusted R^2	0.5796	0.7193	0.7057	0.2081	0.4585	0.4087	0.6519	0.7444	0.7348

Notes: Heteroskedasticity-consistent standard errors clustered at the municipal level are in parenthesis. The unit of observation is the municipality. This study uses the weight of the number of people aged 15 years and above in each municipality to consider different population sizes across municipalities. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and working experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

Table 5 Second-Step Estimation Results of Distinguishing Urban Wage Premium from Human Capital Externalities by Workers' Skills

	Dependent Variable: Municipal Wage Fixed-Effect, $\log\left(\bar{w}_a^{\text{Type}}\right)$								
	Total Workers			High-Skilled Workers			Low-Skilled Workers		
	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(Population Density)	0.0018	0.0057	0.0051	−0.0007	0.0040	0.0017	0.0024	0.0054	0.0047
	(0.0101)	(0.0088)	(0.0085)	(0.0075)	(0.0062)	(0.0062)	(0.0108)	(0.0095)	(0.0092)
Average Years of Schooling	0.1018***	0.0481***	0.0430***	0.0243*	0.0390***	0.0285***	0.0756***	0.0508***	0.0465***
	(0.0138)	(0.0103)	(0.0097)	(0.0137)	(0.0098)	(0.0091)	(0.0128)	(0.0103)	(0.0098)
State Fixed-Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Municipalities	527	527	527	481	481	481	527	527	527
Adjusted R^2	0.5788	0.7198	0.7060	0.2063	0.4578	0.4075	0.6513	0.7448	0.7350

Notes: Heteroskedasticity-consistent standard errors clustered at the municipal level are in parenthesis. The unit of observation is the municipality. This study uses the weight of the number of people aged 15 years and above in each municipality to consider different population sizes across municipalities. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and working experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

Table 6 Robustness Check of the Second-Step Estimation Results

	Dependent Variable: Municipal Wage Fixed-Effect, $\log\left(\bar{w}_a^{\text{Type}}\right)$								
	Total Workers			High-Skilled Workers			Low-Skilled Workers		
	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(Population Density)	−0.0079	−0.0033	−0.0029	−0.0131	−0.0054	−0.0060	−0.0055	−0.0033	−0.0029
	(0.0126)	(0.0104)	(0.0100)	(0.0100)	(0.0082)	(0.0081)	(0.0133)	(0.0112)	(0.0109)
Average Years of Schooling	0.0765***	0.0244***	0.0222***	−0.0087	0.0141	0.0081	0.0550***	0.0282***	0.0266***
	(0.0116)	(0.0079)	(0.0076)	(0.0141)	(0.0088)	(0.0084)	(0.0087)	(0.0079)	(0.0076)
Log(Housing with H.E.A)	0.0299***	0.0279***	0.0245***	0.0388***	0.0293***	0.0240***	0.0244***	0.0267***	0.0235***
	(0.0113)	(0.0074)	(0.0078)	(0.0133)	(0.0087)	(0.0088)	(0.0092)	(0.0078)	(0.0081)
State Fixed-Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Municipalities	527	527	527	481	481	481	527	527	527
Adjusted R^2	0.5857	0.7315	0.7161	0.2211	0.4692	0.4158	0.6573	0.7547	0.7434

Notes: Heteroskedasticity-consistent standard errors clustered at the municipal level are in parenthesis. The unit of observation is the municipality. This study uses the weight of the number of people aged 15 years and above in each municipality to consider different population sizes across municipalities. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and working experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

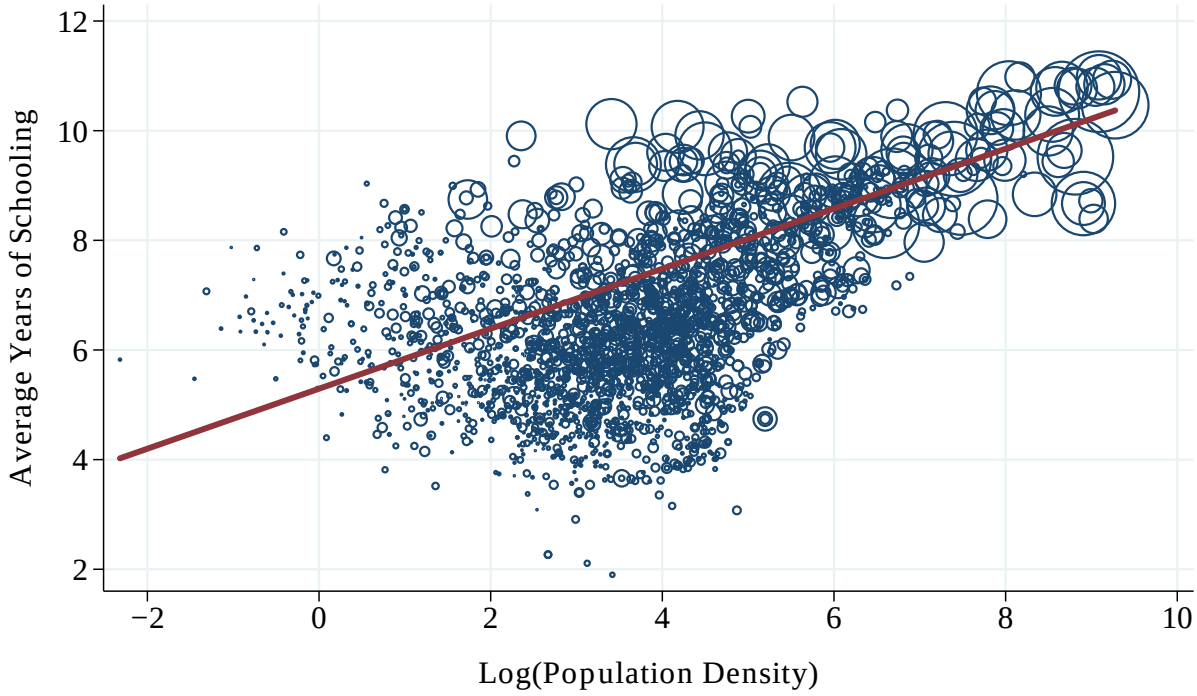
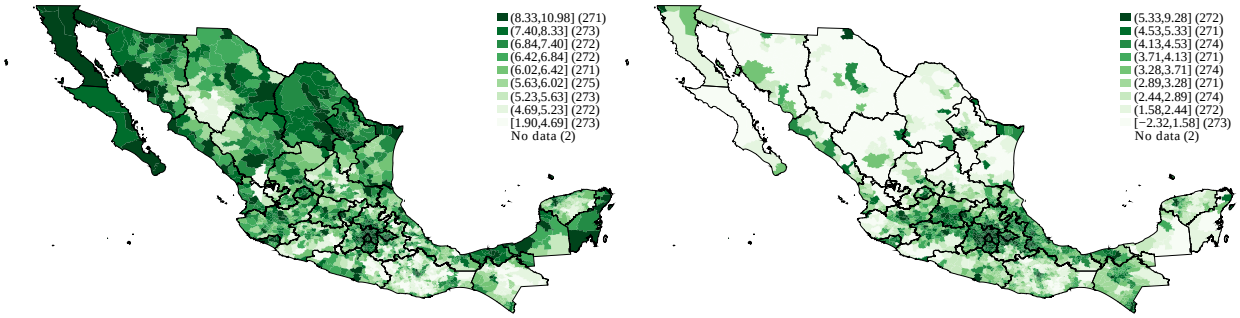


Figure 1 Spatial Sorting of Individuals' Skills

Note: Author's own creation using the 2005 and 2010 Population and Housing Censuses. The average years of schooling are measured for people aged 15 years and above in each municipality. Population density is measured as the total number of people aged 15 years and above divided by the total municipal area (kilometer squared) within a 10 km radius from the centroid of each municipality. The bubble size represents the number of people aged 15 years and above in each municipality. All variables are the average of the 2005 and 2010 Population and Housing Censuses.



(a) Average Years of Schooling **(b)** Population Density

Figure 2 Geographic Distributions of Skilled Individuals

Note: Author's own creation using the 2005 and 2010 Population and Housing Censuses. The average years of schooling are measured for people aged 15 years and above in each municipality. Population density is measured as the total number of people aged 15 years and above divided by the total municipal area (km2) within a 10 km radius from the centroid of each municipality. All variables are the average of the 2005 and 2010 Population and Housing Censuses.

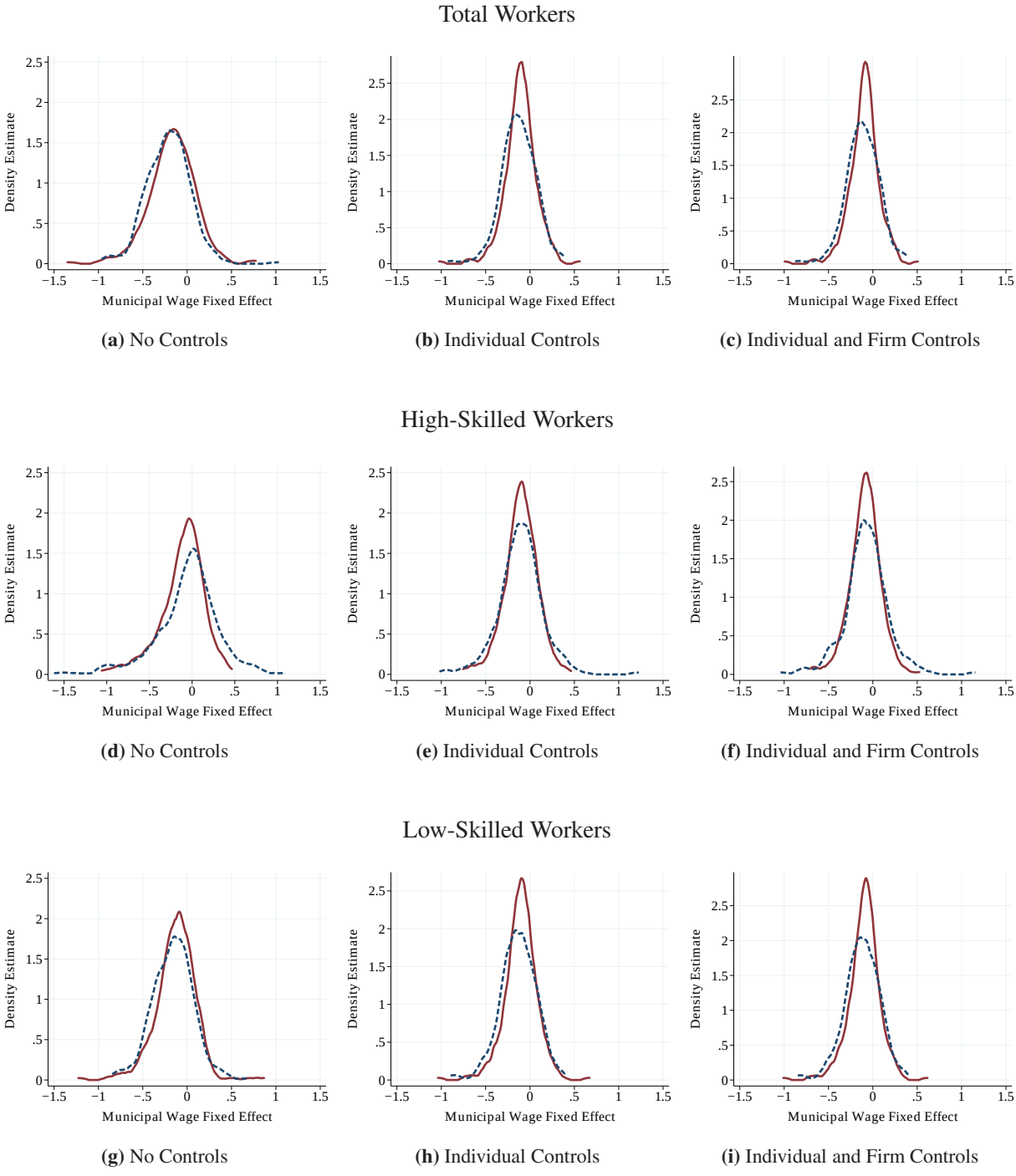


Figure 3 Urban Wage Premium and Spatial Sorting

Notes: The red-solid and blue-dashed lines indicate the municipal wage fixed effect distributions of municipalities with above- and below-median population densities, respectively. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

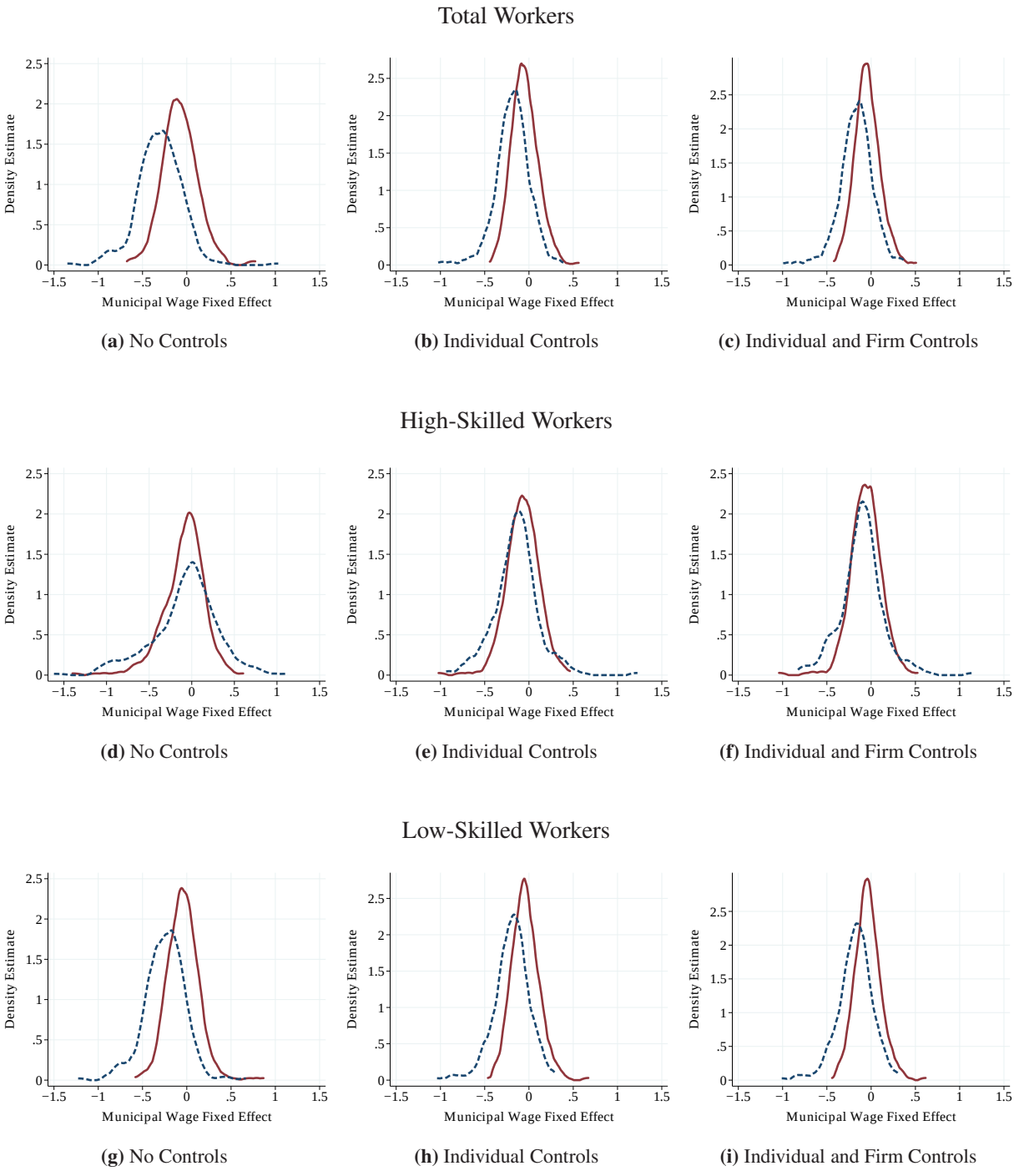
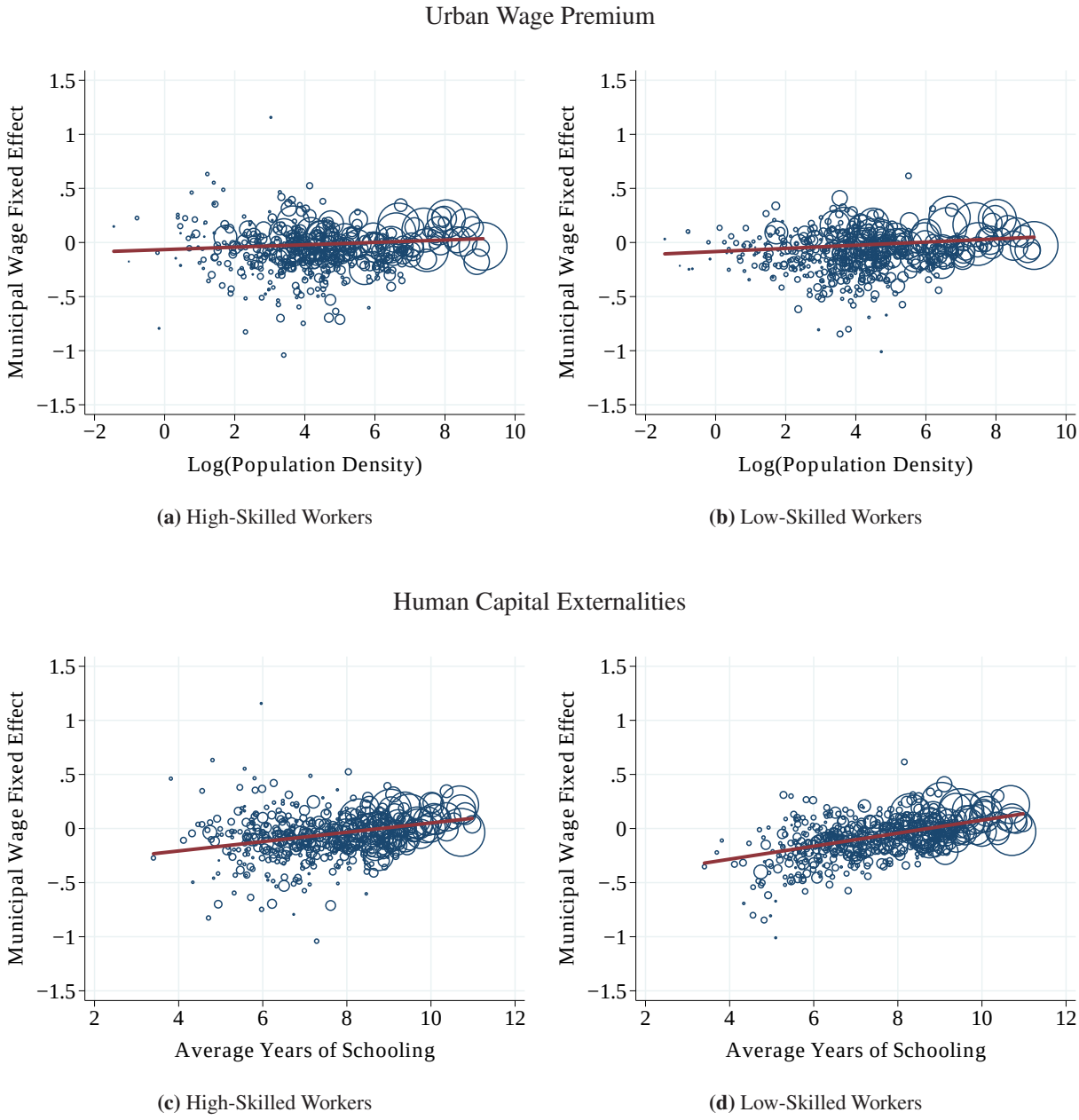


Figure 4 Human Capital Externalities and Spatial Sorting

Notes: The red-solid and blue-dashed lines indicate the municipal wage fixed effect distributions of municipalities with above- and below-median average years of schooling, respectively. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.



(a) High-Skilled Workers

(b) Low-Skilled Workers

(c) High-Skilled Workers

(d) Low-Skilled Workers

Figure 5 Urban Wage Premium and Human Capital Externalities

Note: Author's own creation using the municipal wage fixed effect estimated in Columns (6) and (9) of Table 3. The bubble size represents the number of people aged 15 years and above in each municipality. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

Supplementary Information

Distinguishing the Urban Wage Premium from Human Capital Externalities: Evidence from Mexico

Keisuke Kondo*

RIETI & RIEB, Kobe University

Version: April 11, 2024

Contents

*Corresponding Author. 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).

Appendix A. Controlling for Local Residence and Business Conditions

As a robustness check, this study introduces the number of housing units with home electrical appliances to control residence and business conditions. This section provides additional figures for the number of housing units with home electrical appliances.

Figure A.1 compares the distributions of municipal wage fixed effects between municipalities with above-median (below-median) number of housing units with home electrical appliances by worker skill. In Panels (a) to (c) of Figure A1, both distributions do not overlap by controlling for spatial sorting of worker and ﬁ rm factors, implying that higher wages are paid to workers in better residence and business conditions.

Figure A.2 illustrates local wage premium for high- and low-skilled workers after controlling for spatial sorting. For both workers, this study finds a clear positive relationship between the estimated municipal wage fixed effect and the number of housing units with home electrical appliances, implying that better residence and business locations matter when we consider urban wage premium.

[Figures A.1 and A.2]

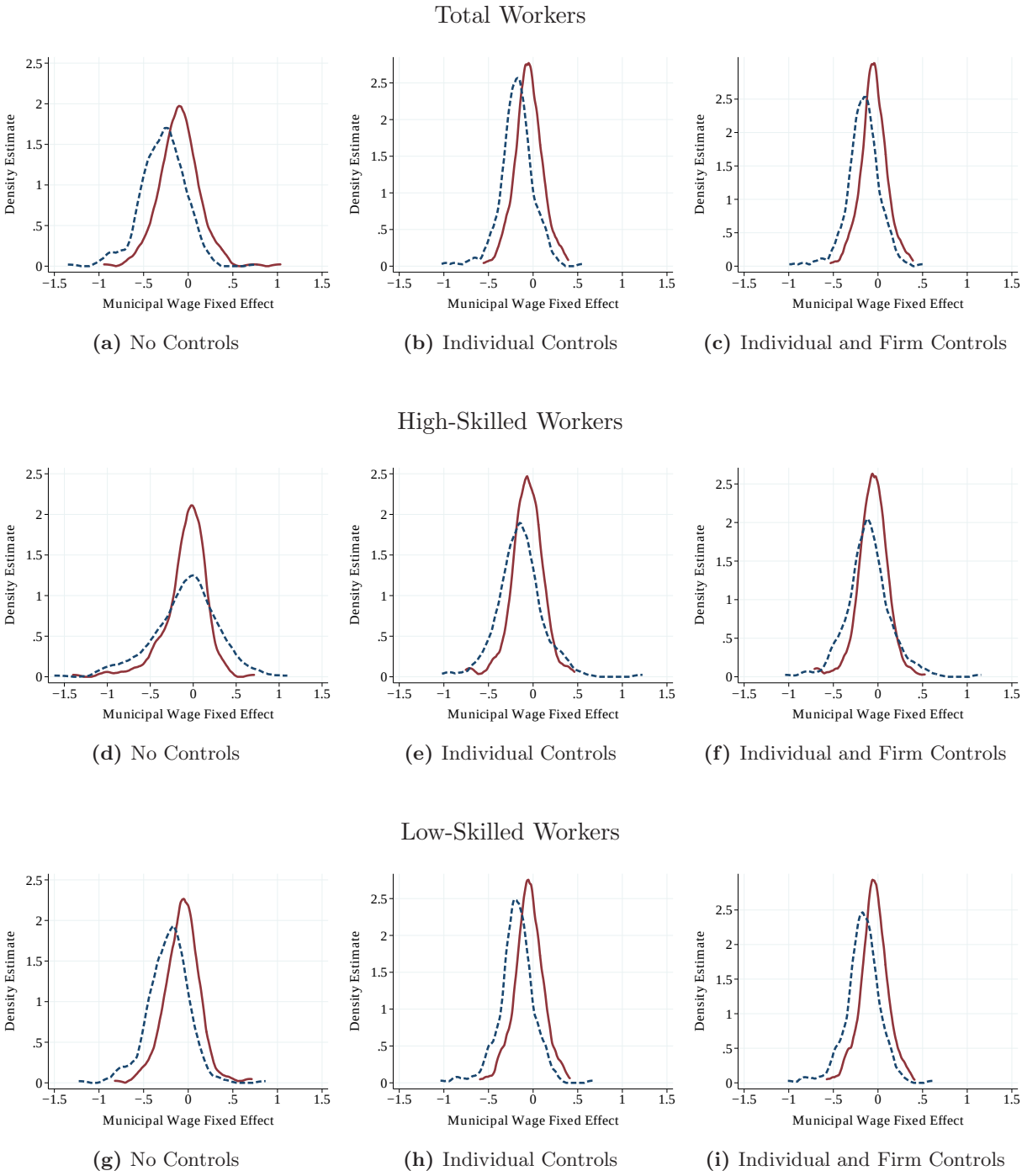
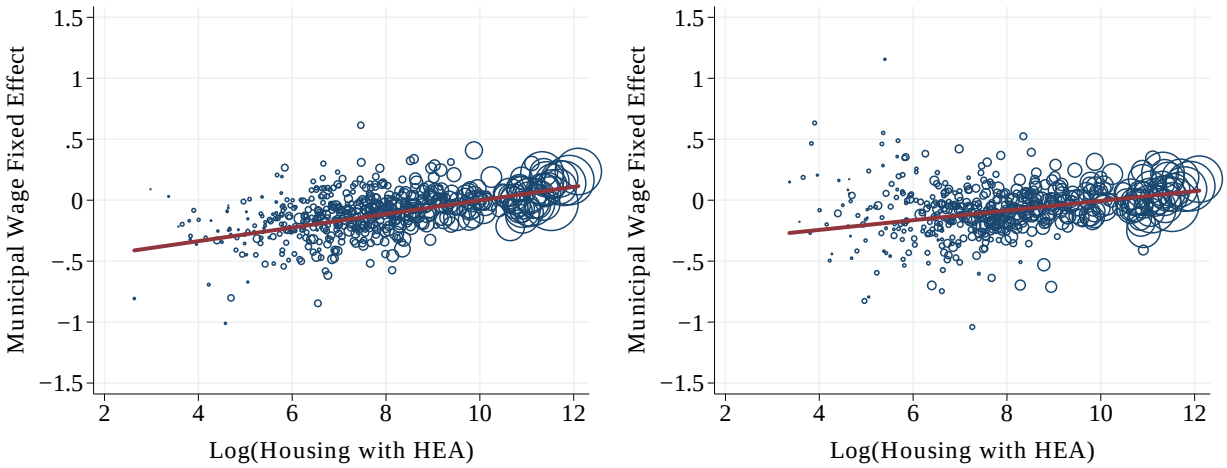


Figure A.1 Location and Spatial Sorting

Notes: The red-solid and blue-dashed lines indicate the municipal wage fixed effect distributions of municipalities with above- and below-median number of housing units with home electrical appliances, respectively. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.



(a) Low-Skilled Workers

(b) High-Skilled Workers

Figure A.2 Wage Premium and Local Residence and Business Condition

Note: Author's own creation using the municipal wage fixed effect estimated in Columns (6) and (9) of Table 3. The bubble size represents the number of people aged 15 years and above in each municipality. High-skilled workers are defined as workers with at least 12 years of schooling in educational institutions, such as vocational schools, colleges, and graduate schools, and experience. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

Appendix B. Population Potential

Although earlier studies uses population/employment density or population size, these variables measured at the administrative unit ignores commuting flows. People cross municipal borders to work and shop. Economic activities are not exclusively conducted within defined administrative areas. Therefore, This study additionally considers population size including the surrounding municipalities.

In the literature of economic geography, Stewart (1947a,b) proposed the notion of "potential of population", inspired by a concept in physics.[1] Following Stewart (1947a,b), the general form of the population potential for municipality c is

$$\mathrm{PP}_a = \sum_{b=1}^{R} P_b D_{ab}^{-\delta}, \tag{B.1}$$

where P_a is the population of municipality a, D_{ab} is the great-circle distance between municipalities a and b, δ is the distance-decay parameter, and R is the number of municipalities. This variable can capture a potential size of population considering the surrounding municipalities. This study uses $\delta = 2$ for the distance decay parameter as a benchmark in the regression analysis.

The regression takes the following form:

$$\log(\bar{w}_a^{\mathrm{Type}}) = \mathrm{Const} + \lambda \log(\mathrm{PP}_a) + \phi \mathrm{Educ}_a + \varepsilon_a, \tag{B.2}$$

where $\log(\bar{w}_a^{\mathrm{Type}}) \in \left\{\log(\bar{w}_a^{\mathrm{N}}), \log(\bar{w}_a^{\mathrm{I}}), \log(\bar{w}_a^{\mathrm{IF}})\right\}$ is the wage premium in municipality a, Const is a constant term, Dens_c is the population density of municipality a, Educ_c is the average years of schooling of municipality a, and ε_a is an error term.

Figure B.1 depicts the geographical distributions of the population potential. The Mexico City metropolitan area shows a higher value of population potential.

Figure B.2 compares the distributions of municipal wage fixed effects between municipalities with above-median (below-median) population potential by worker skill

Figure B.3 illustrates urban wage premium based on the population potential after controlling for spatial sorting.

[1] Stewart (1941) defines "potential of population" as the population divided by distance. The total potentials of population are defined as the summed potentials of population for each region. We call them as "population potential" in this paper.

[Figures B.1–B.3]

Tables B.1–B.3 present the estimation results of the second-step regression by worker skill. Instead of the population density, the population potential is used. We can find that the estimation results of the population potential are almost the same as those of the population density.

[Tables B.1–B.3]

Table B.1 Second-Step Estimation Results of Wage Premium by Worker Skill

	Dependent Variable: Municipal Wage Fixed-Effect, $\log\left(\bar{w}_a^{\text{Type}}\right)$								
	Total Workers			High-Skilled Workers			Low-Skilled Workers		
	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(Population Potential, $\delta = 2$)	0.0644***	0.0350***	0.0313***	0.0120*	0.0254***	0.0169***	0.0478***	0.0357***	0.0325***
	(0.0119)	(0.0076)	(0.0073)	(0.0063)	(0.0054)	(0.0050)	(0.0088)	(0.0079)	(0.0075)
State Fixed-Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Municipalities	527	527	527	481	481	481	527	527	527
Adjusted R^2	0.3971	0.6376	0.6329	0.1948	0.4084	0.3788	0.5175	0.6597	0.6570

Notes: Heteroskedasticity-consistent standard errors clustered at the municipal level are in the parenthesis. The unit of observation is the municipality. This study uses the weight as the number of people aged 15 years and above in each municipality to consider different population sizes across municipalities. High-skilled workers are defined as workers with at least 12 years of schooling and more experiences in educational institutions, such as vocational schools, colleges, and graduate schools. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

Table B.2 Second-Step Estimation Results of Wage Premium by Worker Skill

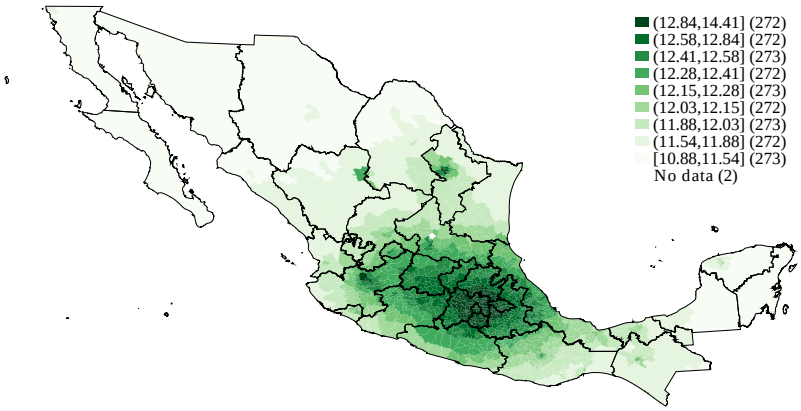
	Dependent Variable: Municipal Wage Fixed-Effect, $\log\left(\bar{w}_a^{\text{Type}}\right)$								
	Total Workers			High-Skilled Workers			Low-Skilled Workers		
	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(Population Potential, $\delta = 2$)	0.0040	0.0049	0.0044	−0.0021	0.0018	−0.0002	0.0021	0.0038	0.0034
	(0.0092)	(0.0093)	(0.0088)	(0.0052)	(0.0055)	(0.0055)	(0.0104)	(0.0101)	(0.0095)
Average Years of Schooling	0.1012***	0.0505***	0.0450***	0.0248**	0.0415***	0.0300***	0.0766***	0.0534***	0.0488***
	(0.0116)	(0.0086)	(0.0082)	(0.0110)	(0.0088)	(0.0082)	(0.0104)	(0.0084)	(0.0081)
State Fixed-Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Municipalities	527	527	527	481	481	481	527	527	527
Adjusted R^2	0.5790	0.7194	0.7057	0.2064	0.4574	0.4074	0.6512	0.7443	0.7346

Notes: Heteroskedasticity-consistent standard errors clustered at the municipal level are in the parenthesis. The unit of observation is the municipality. This study uses the weight as the number of people aged 15 years and above in each municipality to consider different population sizes across municipalities. High-skilled workers are defined as workers with at least 12 years of schooling and more experiences in educational institutions, such as vocational schools, colleges, and graduate schools. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

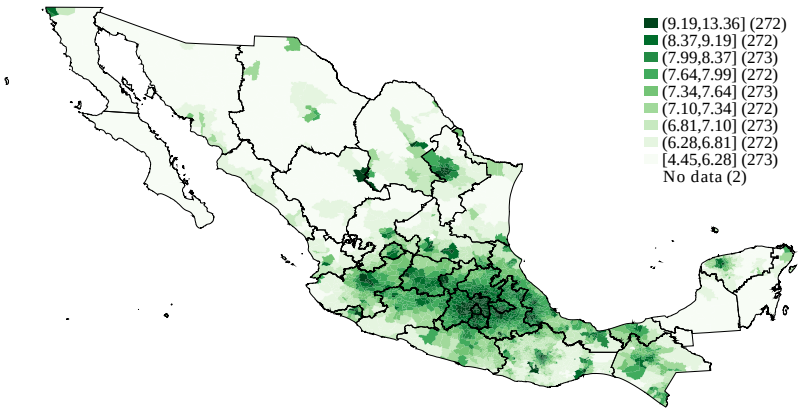
Table B.3 Robustness Check of Second-Step Estimation Results

	Dependent Variable: Municipal Wage Fixed-Effect, $\log\left(\bar{w}_a^{\text{Type}}\right)$								
	Total Workers			High-Skilled Workers			Low-Skilled Workers		
	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls	No controls	Individual Controls	Individual & Firm Controls
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(Population Potential, $\delta = 2$)	0.0015	0.0024	0.0022	−0.0052	−0.0007	−0.0021	0.0000	0.0014	0.0013
	(0.0093)	(0.0090)	(0.0086)	(0.0050)	(0.0055)	(0.0057)	(0.0103)	(0.0097)	(0.0093)
Average Years of Schooling	0.0739***	0.0227***	0.0207***	−0.0101	0.0130	0.0074	0.0536***	0.0268***	0.0254***
	(0.0106)	(0.0076)	(0.0072)	(0.0135)	(0.0087)	(0.0082)	(0.0081)	(0.0077)	(0.0073)
Log(Housing with H.E.A)	0.0253***	0.0257***	0.0226***	0.0323***	0.0264***	0.0210***	0.0214***	0.0247***	0.0217***
	(0.0089)	(0.0062)	(0.0066)	(0.0107)	(0.0076)	(0.0078)	(0.0066)	(0.0064)	(0.0067)
State Fixed-Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Municipalities	527	527	527	481	481	481	527	527	527
Adjusted R^2	0.5848	0.7314	0.7160	0.2184	0.4685	0.4149	0.6567	0.7544	0.7432

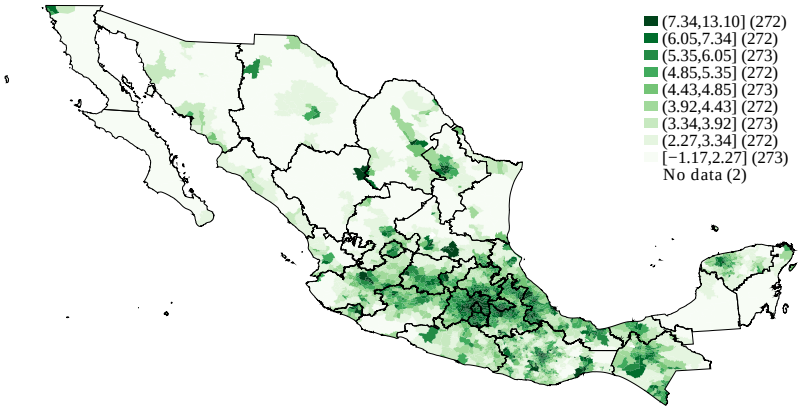
Notes: Heteroskedasticity-consistent standard errors clustered at the municipal level are in the parenthesis. The unit of observation is the municipality. This study uses the weight as the number of people aged 15 years and above in each municipality to consider different population sizes across municipalities. H.E.A. stands for Home Electrical Appliances. High-skilled workers are defined as workers with at least 12 years of schooling and more experiences in educational institutions, such as vocational schools, colleges, and graduate schools. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.



(a) Population Potential ($\delta = 1$)



(b) Population Density ($\delta = 2$)



(c) Population Density ($\delta = 3$)

Figure B.1 Geographic Distributions of Population Potential

Note: Author's creation using the 2005 and 2010 Population and Housing Censuses. Population density is measured the total number of people aged 15 years and above divided by the total municipal area (km^2) within a 10 km radius from the centroid of each municipality. All variables are averaged between the 2005 and 2010 Population and Housing Censuses.

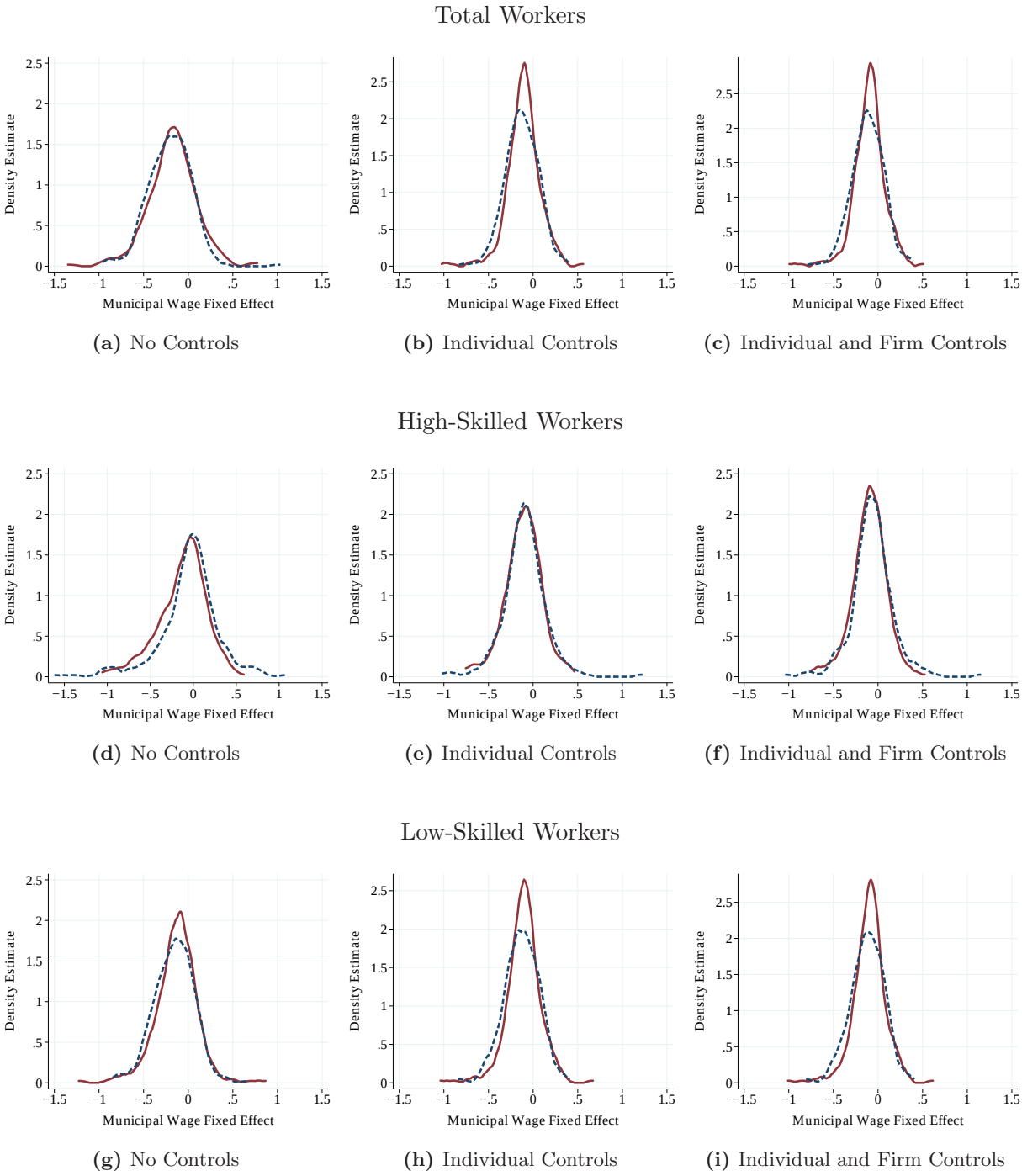
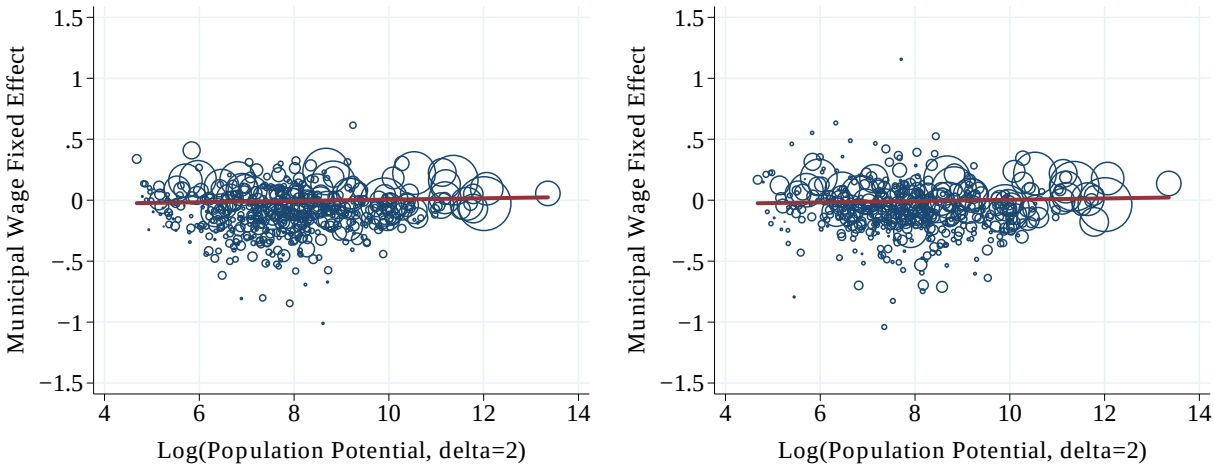


Figure B.2 Population Potential and Spatial Sorting

Notes: The red-solid and blue-dashed lines indicate the municipal wage premium distributions of municipalities with above-median and below-median population density, respectively. High-skilled workers are defined as workers with at least 12 years of schooling and more experiences in educational institutions, such as vocational schools, colleges, and graduate schools. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.



(a) Low-Skilled Workers

(b) High-Skilled Workers

Figure B.3 Wage Premium and Population Potential

Note: Author's creation using the municipal wage premium estimated in Column (6) and (9) of Table 3. The bubble size represents the number of people aged 15 years and above in each municipality. High-skilled workers are defined as workers with at least 12 years of schooling and more experiences in educational institutions, such as vocational schools, colleges, and graduate schools. Low-skilled workers are defined as workers with less than 12 years of schooling and less than a high school diploma.

References

Stewart JQ (1941) An inverse distance variation for certain social influences. Science, New Series 93(2404):89–90. https://doi.org/10.1126/science.93.2404.89

Stewart JQ (1947a) Empirical mathematical rules concerning the distribution and equilibrium of population. Geographical Review 37(3):461–475. https://doi.org/10.2307/211132

Stewart JQ (1947b) Suggested principles of 'social physics'. Science, New Series 106(1748):179–180. https://doi.org/10.1126/science.106.2748.179