

Do Agglomeration Externalities Drive Regional Wages? Evidence from Indonesian Microdata

Icha Wahyu Kusuma Ningrum¹, Ribut Nurul Tri Wahyuni^{2*}, Arya Candra Kusuma³

¹Ministry of Finance, Indonesia

²Politeknik Statistika STIS, Indonesia

³Badan Pusat Statistik, Indonesia

Email: ¹icha.wahyu@kemenkeu.go.id, ²rnurult@stis.ac.id, ³aryacandra@bps.go.id

^{*}Corresponding Author

JEL Classification:

J31

R12

C26

Received: 08 October 2025

Revised: 13 June 2026

Accepted: 18 June 2026

Published online: September 2026

Abstract

Research Originality: This study's novelty lies in using nighttime light (NTL) intensity to measure local density and clay content as an instrumental variable to address endogeneity in Indonesia.

Research Objectives: This study aims to analyze the determinants of regional wage disparities in Indonesia, with particular focus on the role of agglomeration externalities.

Research Methods: This study utilized microdata from BPS-Statistics Indonesia and big data from 2014 to 2024. The analysis employed the instrumental variable two-stage least squares (IV-2SLS) approach to address endogeneity in the local density variable.

Empirical Result: The IV-2SLS estimates indicate that specialization has a positive effect on regional wages, consistent with MAR externality theory, whereas diversity and competition have significant negative effects. Local density exerts a positive effect on regional wages, with the NTL-based estimate yielding a higher elasticity, reflecting its superior capacity to capture actual urban economic density in Indonesia.

Implications: These findings suggest that regional development policies should promote sectoral specialization based on each region's comparative advantage and support urban densification through infrastructure investment and secondary city development to reduce regional wage disparities in Indonesia.

Keywords:

urbanization economies; localization economies; wage disparities; labor characteristics; two-stage least squares

How to Cite:

Ningrum, I. W. K., Wahyuni, R. N. T., & Kusuma, A. C. (2026). Do Agglomeration Externalities Drive Regional Wages? Evidence from Indonesian Microdata. *Etikonomi*, 25(2), 593 – 610. <https://doi.org/10.15408/etk.v25i2.41631>.

INTRODUCTION

Indonesia continues to face various challenges in achieving the Sustainable Development Goals (SDGs) (Putra et al., 2024). These challenges encompass poverty, inequality, infrastructure, energy, environmental sustainability, and employment (Rismaya & Yusuf, 2026). One of the central targets within the SDGs is promoting inclusive and sustainable economic growth alongside productive employment and decent work for all. This objective aligns closely with Indonesia's national development agenda, which emphasizes improving welfare and enhancing the population's quality of life. In this context, the labor sector plays a strategic role in development because labor is a primary factor of production and contributes significantly to economic activity, particularly in developing countries (Umair et al., 2024). Nevertheless, employment-related issues that are closely associated with worker welfare remain substantial challenges in Indonesia.

One important dimension of worker welfare is wages. Wages represent the primary source of income for workers and their households, enabling the fulfillment of basic needs such as food, clothing, housing, education, and health services. According to Parsons & Rabhi (2026), a reduction in wages can decrease household purchasing power and suppress aggregate consumption, ultimately affecting broader economic performance. Consequently, wages are frequently regarded as a key indicator of economic well-being and labor market quality (Ranta & Ylinen, 2024). However, despite sustained economic development, wage disparities between regions in Indonesia remain considerable (Rismaya & Yusuf, 2026).

Wage inequality has become an important concern in labor economics and regional development studies. Discussions regarding regional wage disparities continue to attract the attention of economists and policymakers (Barufi et al., 2023; Ridhwan, 2021). Regional income inequality is observed in almost all countries, although disparities tend to be more pronounced in developing countries. Tabash et al. (2024) argue that income distribution in developing countries is generally less equal than in industrialized countries, implying that developing nations are more vulnerable to persistent regional disparities. As one of the world's largest developing countries, Indonesia faces similar challenges (Rismaya & Yusuf, 2026). The country's vast archipelagic geography, large population, and heterogeneous regional characteristics create substantial variation in economic structure and labor market conditions across regions.

These regional differences are reflected in wage distribution across Indonesia. Average hourly wages vary considerably between regencies and cities. Several regions in Kalimantan and Papua tend to exhibit relatively high wages. In contrast, many regions in Java, Sumatra, and Sulawesi still record comparatively lower wage levels, except in major urban centers such as DKI Jakarta (Maghfirah & Samosir, 2022). This condition indicates that wage inequality across regions remains a substantial issue within the Indonesian labor market (Maghfirah & Samosir, 2022; Ridhwan, 2021; Rismaya & Yusuf, 2026).

Persistent wage disparities may contribute to broader income inequality and uneven welfare distribution. Income inequality itself is associated with numerous adverse

consequences, including lower social welfare, unequal economic opportunities, and various health and social problems (Rismaya & Yusuf, 2026). In Indonesia, these disparities may be more pronounced because regional characteristics differ considerably in geography, infrastructure, industrial concentration, demographic composition, and resource endowment (Ridhwan, 2021). Consequently, understanding the determinants of regional wage differences is essential to support equitable development and improve societal welfare.

Theoretically, wage determination has traditionally been explained through worker characteristics. Rotman & Mandel (2023) emphasized that education, experience, gender, and worker skills constitute the primary determinants of wages. Workers with stronger human capital tend to be more productive, which increases their earnings. Therefore, employers often use worker skills and qualifications to determine wages (Sun, 2026). However, this conventional perspective does not fully account for the role of spatial and regional factors in shaping wage disparities. Economic geography literature argues that regional wage differences may also emerge because of agglomeration externalities generated by the spatial concentration of economic activities (L. Chen et al., 2022; Combes & Gobillon, 2015; Maghfirah & Samosir, 2022).

Agglomeration externalities arise when economic concentration generates productivity advantages through mechanisms such as market pooling, knowledge spillovers, and infrastructure sharing (Ikeda, 2020). Agglomeration externalities are fundamentally categorized into urbanization economies, which arise from overall economic density, and localization economies, which emerge from industrial concentration within a particular sector (Prasertsoong & Puttanapong, 2022). Regions with greater economic density and industrial specialization tend to experience higher productivity (L. Chen et al., 2022) and consequently, higher wages. Previous studies generally classify specialization, diversity, and competition as components of localization economies, while local density represents urbanization economies (Combes & Gobillon, 2015; Groot et al., 2014; Prasertsoong & Puttanapong, 2022). Previous studies have shown that agglomeration economies contribute significantly to regional wage formation (Groot et al., 2014; Prasertsoong & Puttanapong, 2022). Nevertheless, empirical evidence from developing countries (Barufi et al., 2023; A. Chen et al., 2021; L. Chen et al., 2022; Grover et al., 2023; Özgüzel, 2023), particularly Indonesia (Ridhwan, 2021), remains relatively limited. Furthermore, the characteristics of cities and regional economies in developing countries may differ substantially from those in industrialized nations (Grover et al., 2023), suggesting that findings from developed countries may not directly apply to Indonesia (Chauvin et al., 2017). Rapid urbanization, informal labor dominance, and uneven regional development create unique economic structures that require further investigation.

Several studies have examined agglomeration externalities and regional wages in Indonesia (Ridhwan, 2021; Rismaya & Yusuf, 2026). However, important gaps remain unresolved. First, previous studies generally relied on broader regional units, potentially masking substantial heterogeneity at lower administrative levels. Considering Indonesia's highly diverse regional characteristics, analysis at the regency and city level may provide more detailed and accurate insights regarding regional wage disparities.

Second, methodological challenges related to endogeneity remain an important issue in estimating agglomeration effects. The relationship between local density and wages is potentially bidirectional because higher density may increase wages, while higher wages may simultaneously attract additional workers and economic activities (Combes & Gobillon, 2015). Previous studies commonly addressed this issue using historical density variables as instruments (Groot et al., 2014; Ridhwan, 2021). However, the relatively short historical span used in Indonesia may weaken instrument validity. Therefore, this study adopts an alternative geological instrument, namely clay content, following Prasertsoong & Puttanapong (2022), to better address endogeneity concerns. Third, this study also extends the measurement of agglomeration externalities by incorporating nighttime light intensity (NTL) data as a proxy for local density. Population density statistics in developing countries may contain measurement errors because of informal labor mobility and seasonal migration patterns. Previous studies suggest that nighttime light intensity can effectively capture urban economic activities and population distribution (Elvidge et al., 2021; Sangkasem & Puttanapong, 2022; Tan et al., 2018). Therefore, utilizing NTL data may provide a more representative depiction of contemporary urban density in Indonesia.

Based on these considerations, this study aims to analyze the influence of agglomeration externalities on regional wages in Indonesia. This study addresses a critical gap in the literature, as empirical evidence on agglomeration externalities and regional wages in developing countries, particularly at the regency and city level in Indonesia, remains limited. The novelty of this research lies in three aspects: using NTL intensity as an alternative proxy for local density, adopting clay content as a geological instrumental variable to address endogeneity concerns better, and extending agglomeration wage analysis to a finer spatial resolution than previous studies. This study contributes to the literature by providing micro-level empirical evidence on how urbanization and localization economies shape spatial wage disparities across Indonesian regencies and cities, thereby supporting the formulation of more balanced and equitable regional development policies.

METHODS

This study used secondary microdata from the National Labor Force Survey (Sakernas) of BPS-Statistics Indonesia (BPS, 2024) for the August survey periods from 2014 to 2024, excluding 2015, 2016, 2017, 2022, and 2023 due to data availability limitations. We constructed worker-level samples through three filtering stages. First, the sample was restricted to individuals aged 15 years and above (BPS, 2025). Second, the analysis included only workers whose primary employment status was classified as employee, agricultural casual worker, or non-agricultural casual worker. Third, observations with missing values in any variables used in the model were excluded. After applying these criteria, the first-stage estimation utilized a total of 974,760 worker observations pooled across the six survey waves.

The agglomeration externality variables were constructed separately using the full Sakernas sample across all employment statuses while incorporating individual sampling

weights. This approach was intended to better capture labor pooling conditions within the regional economy at the district level. The agglomeration externality variables used in this study were specialization (Prasertsoong & Puttanapong, 2022), diversity (Ikeda, 2020), competition (Ridhwan, 2021), and local density (Elvidge et al., 2021; Sangkasem & Puttanapong, 2022), which were constructed following previous studies. In particular, NTL data were employed as a proxy for local density and were derived from Google Earth Engine (Earth Observation Group, 2026).

In addition to Sakernas, this study also utilized secondary data from the Village Potential Survey (Podes) to represent local amenities, particularly the number of restaurants and hotels at the district level. Information regarding natural amenities was obtained from the World Climate (Fick & Hijmans, 2017) and the Indonesia Geospatial Portal (Geospatial Information Agency, 2026). Furthermore, clay content data were obtained from Google Earth Engine. These datasets were subsequently integrated to construct the variables used in the second stage of estimation, resulting in a pooled sample of 8,784 region-sector-year observations during the study period.

To examine the determinants of regional wage disparities, this study employed a two-stage estimation approach following Barufi et al. (2023); Groot et al. (2014); Prasertsoong & Puttanapong (2022); and Ridhwan (2021). To assess the reliability of the estimates, following Catherine et al. (2022), this study ran robustness tests at both estimation stages by systematically adding and removing control variables; consistent coefficient signs and significance levels across specifications confirm the reliability of the estimated results. The first-stage estimation isolated the regional wage component by controlling for individual worker characteristics. Subsequently, the second-stage estimation examined the influence of agglomeration externalities on the regional wage component derived from the first-stage estimation. The empirical first-stage model used in this study is presented in Equation (1).

$$\begin{aligned} \ln W_{irt} = & \beta_0 + \beta_1 Age_{irt} + \beta_2 Age_{irt}^2 + \beta_3 Male_{irt} + \beta_4 Married_{irt} \\ & + \beta_5 Edu_{irt} + \beta_6 WCollar_{irt} + \sum_{r=1}^{N_r-1} \gamma_r D_{irt}^r + \sum_{s=1}^{N_s-1} \mu_s D_{irt}^s \\ & + \sum_{t=1}^{N_t-1} \delta_t D_{irt}^t + \varepsilon_{irt} \end{aligned} \quad (1)$$

Where: W_{irt} represents hourly wage; Age_{irt} represents the worker's age; Age_{irt}^2 represents the worker's age squared to account for non-linear relationships; $Male_{irt}$ represents the worker's gender where 1 is male, and 0 is female; $Married_{irt}$ represents a worker's marital status where 1 is married and 0 otherwise; Edu_{irt} represents a worker's years of schooling; $WCollar_{irt}$ is a dummy variable that represents the worker's type of employment where 1 is white-collar, and 0 is blue-collar; γ_r denotes regional fixed effects; μ_s represents sector fixed effects; δ_t represents year fixed effects; and ε_{irt} represents *i.i.d* errors. i, s, r , and t represent the individual, sector, region, and year, respectively.

Equation (1) represents the application of the ordinary least squares (OLS) approach to examine the impact of worker attributes on wages. When the homoscedasticity

assumption was not met, we applied robust standard errors to ensure valid statistical inference. In assessing the impact of agglomeration variables in the second stage, this study recalibrated the fixed effects in Equation (1) to incorporate sector, region, and year fixed effects (Prasertsoong & Puttanapong, 2022; Ridhwan, 2021). This study regressed the aggregated fixed effects on the agglomeration externality metrics, temporal fixed effects, and sectoral fixed effects. The combined region-sector-year fixed effects obtained from the first-stage regression constituted the spatial residual (γ_{srt}), the regional wage component that remained after correcting for worker characteristics and industrial composition, which served as the dependent variable in the second-stage. The second-stage estimation model can be written as in Equation (2).

$$\begin{aligned} \gamma_{srt} = & \beta_0 + \beta_1 Specialization_{srt} + \beta_2 Diversity_{rt} \\ & + \beta_3 Competition_{rt} + \beta_4 \ln Density_{rt} + \beta_5 \ln Hore_{rt} \\ & + \beta_6 \ln Dist_{coas_r} + \beta_7 \ln Dist_{prov_r} + \beta_8 Temp_r \\ & + \beta_9 Rainfall_r + \beta_{10} Elev_r + \beta_{11} Area_r + \mu_s + \delta_t \\ & + error_{srt} \end{aligned} \quad (2)$$

Where: $Specialization_{srt}$ represents specialization (Prasertsoong & Puttanapong, 2022); $Diversity_{rt}$ represents diversity (Ikeda, 2020); $Competition_{rt}$ represents competition (Ridhwan, 2021), $Density_{rt}$ represents local density (Elvidge et al., 2021; Sangkasem & Puttanapong, 2022), $Hore_{rt}$ represents the number of hotels-restaurants (Ridhwan, 2021); $Dist_{coas_r}$ represents the distance of district center from the nearest coastline (Beltrán Tapia et al., 2018); $Dist_{prov_r}$ represents the distance between the district center and the provincial capital center (Beltrán Tapia et al., 2018); $Temp_r$ represents temperature (Chauvin et al., 2017); $Rainfall_r$ represents rainfall (Chauvin et al., 2017); $Elev_r$ represents elevation (Yang et al., 2022); $Area_r$ represents district area; μ_s represents sector fixed effects; δ_t represents year fixed effects.

To examine the impact of agglomeration externalities on regional wages, this study employed the instrumental variable two-stage least squares (IV2SLS) method to address the simultaneity bias arising from the endogenous relationship between local density and regional wages (Prasertsoong & Puttanapong, 2022). Since local density and wages can be simultaneously determined, estimating this relationship using OLS would yield inconsistent coefficient estimates. To resolve this, population density is treated as the endogenous explanatory variable, and the percentage of clay content in soil is used as the instrumental variable. Following Prasertsoong & Puttanapong (2022), clay content satisfies both conditions required of a valid instrument: it is positively correlated with current population density through the persistence of historical settlement patterns driven by rice cultivation on clay-rich soils, and it does not directly affect current regional wages because contemporary urban economic activity is dominated by service and industry-based sectors in which soil fertility is no longer relevant. In addition, equation (2) was also estimated by substituting population density with nighttime light (NTL) intensity as an alternative proxy for local density, following Prasertsoong & Puttanapong (2022).

To verify the validity of the IV2SLS estimation, we conducted three diagnostic tests. First, an endogeneity test using the GMM distance statistic was performed to confirm whether population density is statistically endogenous; rejection of the null hypothesis of exogeneity validates the use of IV2SLS over OLS. Second, an underidentification test using the Kleibergen-Paap rk Lagrange multiplier statistic (Kleibergen & Paap, 2006) was performed to test whether the instrument is correlated with the endogenous variable and whether the equation is correctly identified. Third, a weak identification test using the Kleibergen-Paap rk Wald F-statistic was conducted to evaluate the instrument's predictive power; a value exceeding the threshold of 10 confirms that the instrument is not weak. Robust standard errors were applied throughout to address heteroscedasticity.

RESULTS AND DISCUSSION

This study reveals three main findings. First, worker characteristics, including age, gender, marital status, years of schooling, and occupational type, collectively explain 90.98% of the variation in individual hourly wages, with all variables exhibiting statistically significant effects in directions consistent with human capital theory. Second, among localization economy externalities, specialization exerts a positive and statistically significant effect on regional wages, consistent with the Marshall-Arrow-Romer (MAR) externality theory. In contrast, diversity and competition each exert negative and significant effects, a pattern previously corroborated by evidence from the Netherlands (Groot et al., 2014) and Thailand (Prasertsoong & Puttanapong, 2022), and reported here for the first time at the regency and city level in Indonesia. Third, local density, as a proxy for urban economies, exerts a positive and significant effect on regional wages, with an elasticity of 0.0442 when proxied by population density and 0.0528 when proxied by nighttime light (NTL) intensity. The higher estimate using NTL may reflect the superior ability of satellite-derived luminosity data to capture actual urban economic density in Indonesia. Collectively, these findings demonstrate that regional wage disparities in Indonesia are shaped not only by individual worker attributes but also by the agglomeration structure of local economies.

Table 1 presents the descriptive statistics of worker-level variables used in the first-stage estimation, pooled across six survey waves from 2014 to 2024. Overall, mean hourly wages trend upward throughout the study period, except in 2020, when they declined due to the economic contraction induced by the COVID-19 pandemic (Kemenko Perekonomian, 2021). Educational attainment also improves gradually over time, though the average years of schooling remains below the 12-year universal secondary education threshold, suggesting that a substantial proportion of the Indonesian workforce has yet to complete formal secondary schooling. In terms of sample composition, the data are dominated by male, married, and blue-collar workers, with the tertiary sector accounting for the largest share of employment across all survey years. These compositional characteristics motivate the inclusion of gender, marital status, occupational type, and sector dummies as control variables in the first-stage regression.

Table 1. Descriptive Statistics of the Variables in the First-Stage Estimation

Variables	2014	2018	2019	2020	2021	2024	2014–2024
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln (Hourly Wage)	8.96 (0.87)	9.44 (0.80)	9.48 (0.78)	9.39 (0.84)	9.40 (0.84)	14.59 (0.81)	10.45 (2.29)
Age	37.21 (11.89)	37.33 (11.72)	37.21 (11.79)	38.46 (12.33)	38.63 (12.29)	38.74 (12.32)	38.06 (12.13)
Duration of Schooling	9.91 (4.73)	11.15 (4.35)	11.16 (4.33)	10.20 (4.41)	10.30 (4.52)	10.77 (4.43)	10.57 (4.48)
Male (base: Female)	69.49%	65.42%	65.83%	67.55%	67.10%	65.30%	66.70%
Married (base: Unmarried)	76.11%	75.15%	74.77%	77.12%	76.71%	75.35%	75.93%
White-collar (base: Blue-collar)	30.66%	39.23%	38.75%	29.61%	30.33%	32.92%	33.16%
Tertiary Sector	53.81%	64.53%	63.69%	53.32%	53.45%	57.43%	57.23%
Observations	119,247	104,662	166,129	189,530	191,356	203,836	974,760

Note: This study reports continuous variables as means and standard deviations (in parentheses), and categorical variables as the percentage of workers in each category

Table 2. Descriptive Statistics of Agglomeration Variables

Variables	2014	2018	2019	2020	2021	2024	2014–2024
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Specialization	0.34 (0.23)	0.34 (0.22)	0.34 (0.22)	0.34 (0.22)	0.34 (0.21)	0.34 (0.23)	0.34 (0.22)
Diversity	0.85 (0.18)	0.86 (0.18)	0.88 (0.17)	0.87 (0.17)	0.89 (0.17)	0.85 (0.19)	0.87 (0.18)
Competition	0.99 (0.00)	0.99 (0.00)	0.99 (0.00)	0.99 (0.00)	0.99 (0.00)	0.99 (0.00)	0.99 (0.00)
Ln Population Density	5.29 (1.93)	5.35 (1.91)	5.33 (1.92)	5.37 (1.90)	5.40 (1.90)	5.41 (1.89)	5.36 (1.91)
Ln NTL	0.59 (0.95)	0.72 (1.03)	0.73 (1.04)	0.75 (1.03)	0.78 (1.05)	0.87 (1.11)	0.74 (1.04)
Ln Hore	4.36 (1.47)	4.65 (1.47)	4.99 (1.58)	4.93 (1.52)	4.98 (1.46)	5.14 (1.48)	4.84 (1.52)
Observations	1,468	1,440	1,466	1,469	1,469	1,472	8,784

Note: This study reports continuous variables as means and standard deviations (in parentheses)

Table 2 presents descriptive statistics for the agglomeration variables used in the second-stage estimation, covering 8,784 region-sector-year observations pooled across six survey waves. The agglomeration externality variables, namely specialization, diversity, and competition, remain remarkably stable over the study period, reflecting the structural persistence of agglomeration patterns across Indonesian districts. Of particular note is the contrasting temporal behavior of the two local density proxies:

while population density exhibits only a modest upward trend, nighttime light (NTL) intensity demonstrates a more pronounced and consistent increase over the same period. This divergence suggests that NTL intensity more sensitively captures the dynamics of urban economic expansion compared to population density, which may be subject to measurement limitations arising from informal labor mobility and seasonal migration patterns in Indonesia.

Table 3 presents the first-stage regression results across four model specifications. The simultaneous test results indicate that the p-value is less than the significance level, confirming that worker characteristics collectively exert a statistically significant influence on hourly wages. The robustness check shows that the sign and significance of all coefficients remain consistent across Model 1 through Model 4, regardless of the addition or removal of sector, region, and year dummies, suggesting stable estimation results. This study focuses on Model 4, which constitutes the most complete specification. The adjusted R^2 of Model 4 is 0.9098, indicating that worker characteristics account for 90.98% of the variation in individual wages, with the remaining variation attributable to factors outside individual attributes, including regional characteristics that this study examines in the second stage.

Table 3. Estimation results for the pooled cross-section from 2014–2024

Variables	Model 1	Model 2	Model 3	Model 4
(1)	(2)	(3)	(4)	(5)
Age	0.0595*** (0.0012)	0.0591*** (0.0012)	0.0553*** (0.0012)	0.0402*** (0.0004)
Age Squared	-0.0004*** (0.0000)	-0.0004*** (0.0000)	-0.0003*** (0.0000)	-0.0003*** (0.0000)
Male	0.2363*** (0.0049)	0.2196*** (0.0050)	0.1748*** (0.0050)	0.2843*** (0.0016)
Married	-0.0860*** (0.0073)	-0.0955*** (0.0073)	-0.1070*** (0.0073)	0.1188*** (0.0022)
Duration of Schooling	0.0870*** (0.0006)	0.0929*** (0.0006)	0.0912*** (0.0006)	0.0585*** (0.0002)
White-collar	0.0588*** (0.0062)	0.1049*** (0.0065)	0.0656*** (0.0066)	0.3025*** (0.0020)
Sector Dummies	No	Yes	Yes	Yes
District Dummies	No	No	Yes	Yes
Year Dummies	No	No	No	Yes
Constant	7.8726*** (0.0230)	7.9710*** (0.0233)	7.6950*** (0.0637)	6.7680*** (0.0229)
Prob(F-Stat)	0.0000	0.0000	0.0000	0.0000
Adj. R^2	0.0395	0.0404	0.0544	0.9098

Note: Robust standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

All worker characteristic variables exert statistically significant effects. Age exerts a positive but diminishing effect on wages, reflecting the accumulation of work experience over time, consistent with the findings of Groot et al. (2014), Prasertsoong & Puttanapong (2022), and Ridhwan (2021). Male workers earn considerably higher wages than their female counterparts, reflecting persistent labor market segmentation in Indonesia where female workers remain disproportionately concentrated in lower-wage sectors and informal employment arrangements (Akbar, 2022). The effect of marital status turns positive and statistically significant in Model 4, the preferred specification given its substantially higher adjusted R^2 . The negative coefficients in Models 1 through 3 reflect a temporal composition bias that is absorbed once year fixed effects are introduced, allowing the within-year marriage premium to emerge, consistent with the male breadwinner norm embedded in Indonesian household culture. Each additional year of schooling is associated with a significant wage increase, consistent with human capital theory and reflecting education's role as a productivity-enhancing investment in the Indonesian labor market. Finally, white-collar workers earn significantly higher wages than blue-collar workers, reflecting the premium associated with skill-intensive occupations in a labor market where low-skilled manual employment remains dominant.

Figure 1. Average spatial residual by regency and city, 2014–2024

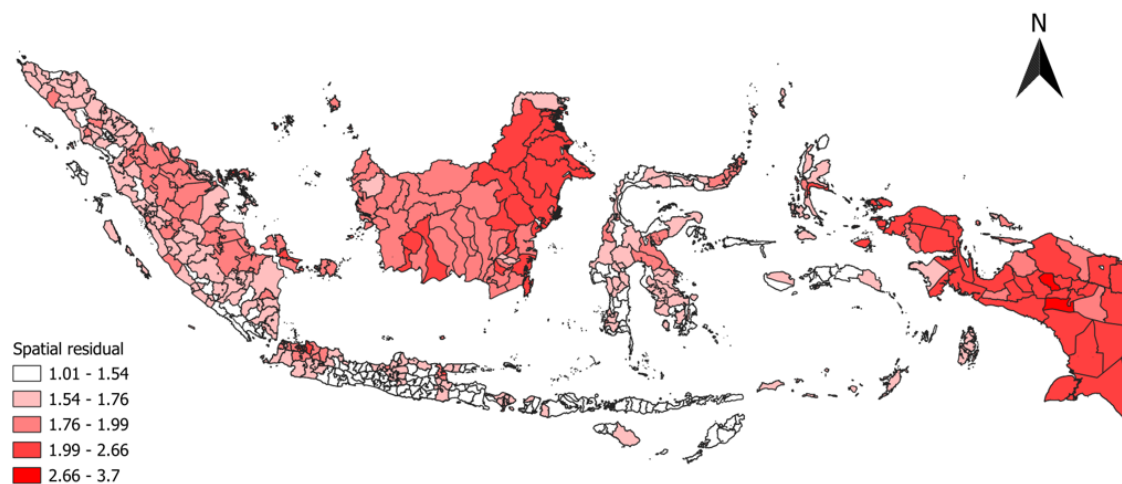


Figure 1 displays the spatial distribution of the regional wage component net of individual worker characteristics across Indonesian regencies and cities, revealing pronounced spatial polarization that cannot be attributed to worker attributes alone. DKI Jakarta exhibits the highest regional wage levels, reflecting its dominance as Indonesia's primary agglomeration center, consistent with evidence from Amsterdam and Bangkok (Groot et al., 2014; Prasertsoong & Puttanapong, 2022). Elevated wages in Kalimantan and Papua, by contrast, are driven not by agglomeration but by compensating wage differentials associated with the extractive primary sector and higher local cost of living. This spatial heterogeneity motivates the second-stage analysis examining the role of agglomeration externalities in explaining regional wage differences.

Table 4. Results of the second-stage estimation regression with population density and NTL intensity as a proxy for local density

Variables	Population density				NTL			
	OLS		IV2SLS		OLS		IV2SLS	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Specialization	0.166*** (0.0202)	0.179*** (0.0193)	0.1761*** (0.0200)	0.177*** (0.0197)	0.176*** (0.0196)	0.178*** (0.019)	0.177*** (0.0196)	0.179*** (0.0191)
Diversity	-0.341*** (0.0219)	-0.272*** (0.0227)	-0.297*** (0.0245)	-0.156*** (0.0282)	-0.250*** (0.0238)	-0.161*** (0.0242)	-0.2395*** (0.0344)	-0.118*** (0.0351)
Competition	-438.4*** (26.623)	-362.42*** (26.635)	-201.47*** (49.8468)	-200.5*** (31.727)	-220.0*** (23.201)	-231.0*** (25.599)	-198.6*** (49.506)	-190.6*** (32.617)
Local Density	-0.012*** (0.0024)	-0.0454*** (0.0027)	0.0365*** (0.0103)	0.044*** (0.0128)	0.047*** (0.004)	0.029*** (0.0049)	0.054*** (0.0148)	0.053*** (0.0150)
Ln Hore		0.0123*** (0.0032)		-0.032*** (0.0069)		-0.016*** (0.0030)		-0.021*** (0.0042)
Ln Dist_coas		-0.006** (0.0024)		-0.009*** (0.0027)		-0.0068** (0.0025)		-0.006** (0.0026)
Ln Dist_prov		-0.033*** (0.0022)		-0.020*** (0.0028)		-0.0225*** (0.0022)		-0.0192*** (0.0029)
Temperature		0.063*** (0.0035)		0.049*** (0.0042)		0.053*** (0.0036)		0.0496*** (0.0041)
Rainfall		0.001*** (0.0001)		0.001*** (0.0001)		0.0011*** (0.0001)		0.0012*** (0.0001)
Elevation		-0.0003*** (0.0000)		-0.0002*** (0.0000)		-0.0002*** (0.0000)		-0.0002*** (0.0000)
Area	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	438.98*** (26.629)	363.09*** (26.643)	201.74*** (49.910)	200.67*** (31.778)	220.41*** (23.206)	231.32*** (25.608)	199.03*** (49.536)	190.88*** (32.646)
Adj R Square	0.9798	0.9821	0.9788	0.9801	0.9800	0.9815	0.9801	0.9616
Prob F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Kleibergen- Paap rk Wald F statistic:			757.92	835.96			655.129	1352.14
Under- identification			0.0000	0.0000			0.0000	0.0000
Endogeneity			0.0000	0.0000			0.6439	0.0844

Note: Robust standard errors are in parentheses; *p<0.1,**p<0.05,***p<0.01

Table 4 presents the second-stage estimation results using both population density and nighttime light (NTL) intensity as proxies for local density. The robustness check confirms that coefficient signs and significance levels remain consistent across Model 1 and Model 2 for both proxies. The sign reversal of the population density coefficient from negative under OLS to positive under IV2SLS demonstrates that OLS produces

biased estimates by failing to account for the simultaneity between density and wages, whereby high-wage regions simultaneously attract more workers and economic activity. The clay content instrument satisfies both the underidentification test (p -value < 0.05) and the weak identification test (Kleibergen-Paap rk Wald F-statistic of 835.96, well above the threshold of 10), confirming its validity (Hu et al., 2026).

Alongside the underidentification and weak identification tests, a monotonicity graph was used to verify that clay content percentage, as the instrument variable, exhibits a positive relationship with local density, consistent with theoretical expectations (Sangkasem & Puttanapong, 2022). As shown in Figure 2, both proxies of the endogenous explanatory variable (population density and NTL intensity) display a positive association with clay content percentage. This confirms that the instrument satisfies the monotonicity condition, whereby clay content maintains a consistent positive association with each proxy for local density across the distribution (Small et al., 2017).

Specialization exerts a positive and statistically significant effect on regional wages, consistent with MAR externality theory, which posits that sectoral concentration facilitates intra-industry knowledge spillovers, labor pooling, and productivity gains (Capello, 2014). In the Indonesian context, specialization premiums are particularly pronounced in two distinct spatial configurations. First, resource-rich regencies outside Java, particularly in Sumatra, Kalimantan, and Papua, benefit from specialization in mining and plantation agriculture, where sector-specific knowledge accumulation, labor pooling in extractive industries, and infrastructure investment collectively sustain elevated wages. Second, Java's manufacturing belt exhibits high secondary sector specialization, where long-standing industrial agglomeration, well-developed supply chain networks, and proximity to input and output markets reinforce self-sustaining productivity advantages. This dual specialization pattern suggests that Indonesia's regional wage premiums are not confined to a single sector but emerge from distinct localization economy mechanisms operating across different regional contexts.

These spatial wage differentials likely reflect the degree to which each region has developed deep sector-specific ecosystems, where repeated interaction among firms, workers, and institutions within the same industry generates compounding returns that generalist economies cannot replicate. This finding corroborates evidence from the Netherlands (Groot et al., 2014) and Thailand (Prasertsoong & Puttanapong, 2022), and extends the positive specialization-wage relationship to the Indonesian regency and city level. As shown in Figure 3, specialization premiums are particularly pronounced in two distinct spatial configurations: primary sector specialization concentrated in resource-rich regencies outside Java, driven by mining and plantation agriculture in Sumatra, Kalimantan, and Papua, and high secondary sector specialization in Java's manufacturing belt, consistent with Chumaidiyah et al. (2023). Across these varied national contexts, the convergent result that specialization raises wages reinforces the robustness of localization economy theory beyond high-income settings and confirms its applicability in archipelagic, resource-dependent economies such as Indonesia. From a policy perspective, these findings support expanding targeted industrial clustering strategies that leverage each region's comparative advantage.

Strengthening Indonesia's existing special economic zone framework, particularly by extending it to resource-rich regions outside Java, could amplify localization economies, deepen sectoral specialization, and raise regional wages in areas currently underserved by agglomeration-driven development.

Figure 2. Monotonicity graph of (a) population density and (b) NTL intensity as a proxy for local density and percentage of clay content

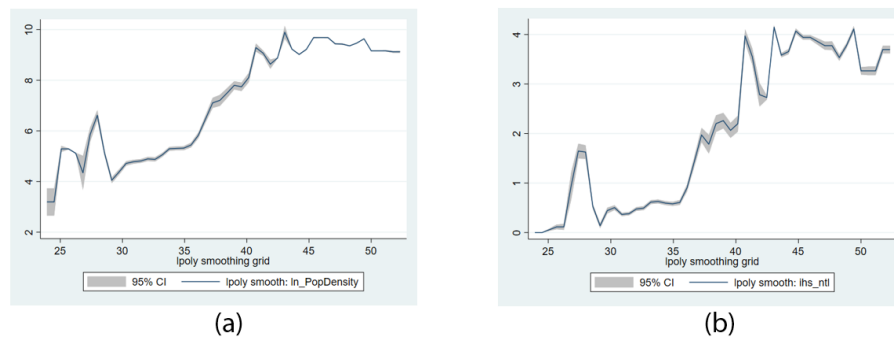
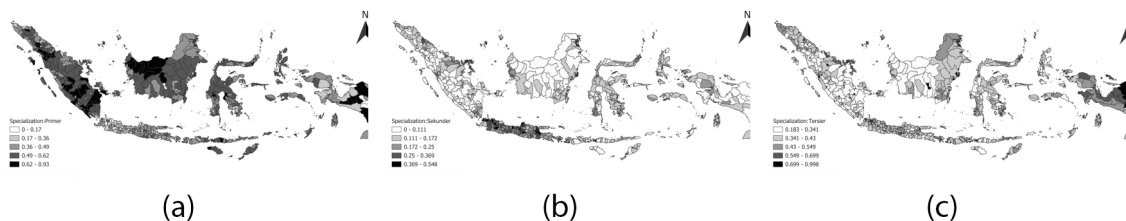


Figure 3. The average agglomeration externalities from 2014 to 2024, broken down by region, were as follows: (a) Specialization: primary; (b) Specialization: secondary; (c) Specialization: tertiary



Diversity has a negative and significant effect on regional wages, which contradicts Jacobs' spillover theory (Ikeda, 2020). In Indonesia, this may reflect that economic diversification at the regency level often involves spreading labor across low-productivity informal activities, such as petty trade, subsistence agriculture, and household services, rather than building productive links across sectors that promote innovation and wage growth. The scale needed for Jacobs externalities may exist only in major metropolitan areas such as DKI Jakarta, Surabaya, and Medan, while most regencies lack the economic scale needed to benefit from diversification. Thus, premature diversification may weaken wages by spreading workers across low-wage activities instead of concentrating them in more productive sectors.

This finding is consistent with evidence from the Netherlands (Groot et al., 2014) and Thailand (Prasertsoong & Puttanapong, 2022), and with Grover et al. (2023), who find that Jacobs externalities are generally weaker in areas with high labor informality. Cross-sector knowledge spillovers require sufficient urban scale and economic development to emerge (Combes & Gobillon, 2015). Therefore, diversity-enhancing policies should be carefully sequenced. Regions should first strengthen their productive base and specialization

before pursuing broader diversification. At the same time, labor formalization, skills development, and knowledge-intensive services can help make diversification more productive.

Competition also has a negative and significant effect on regional wages, contrary to Porter's spillover theory, which argues that competition encourages innovation, productivity, and higher wages (Capello, 2014). In most Indonesian regencies, competition is largely among small informal firms that compete through lower costs rather than innovation. Greater competition may therefore increase pressure to reduce labor costs, limiting wage and productivity growth. Together with the negative diversity effect and positive specialization effect, this result suggests that MAR-type externalities may be more dominant in Indonesia, where informality remains widespread.

This finding is consistent with evidence from the Netherlands (Groot et al., 2014) and Thailand (Prasertsoong & Puttanapong, 2022). In these settings, competition effects can be negative when informal micro-enterprises dominate and lack the capacity to convert competitive pressure into innovation. To strengthen the benefits of competition, Indonesia should reduce barriers to firm formalization and promote quality-based competition. Policies should also support innovation through R&D investment, technology adoption, and stronger links between industry and higher education institutions.

Local density has a positive and significant effect on regional wages under the IV-2SLS estimation, confirming the presence of urbanization economies across Indonesian regencies and cities. The change from a negative coefficient under OLS to a positive coefficient under IV-2SLS highlights the importance of addressing endogeneity. High-wage regions may attract more workers and economic activity, which can bias OLS estimates downward. The clay content instrument satisfies the relevance and exclusion conditions, supporting the validity of the IV approach. The elasticity of NTL intensity is higher than that of population density, suggesting that nighttime light better captures actual urban economic activity. This is particularly relevant in Indonesia, where informal employment and seasonal labor mobility may cause official population data to understate the economically active population. The faster growth of NTL intensity compared with population density also indicates that urban economic activity is expanding faster than population figures suggest.

These findings are consistent with evidence from Brazil (Barufi et al., 2023), Turkey (Özgüzel, 2023), and Thailand (Prasertsoong & Puttanapong, 2022). The advantage of NTL as a density measure also supports Sangkasem & Puttanapong (2022) and Elvidge et al. (2021), while extending their findings to the Indonesian regency and city level. The findings have two key policy implications. First, investment in urban infrastructure, including transport, digital connectivity, and public services, should strengthen existing agglomeration centers. Second, developing secondary cities outside Java through infrastructure investment, decentralization, and incentives for private-sector relocation could create new growth centers, reduce Java-centric development, and narrow regional wage disparities.

CONCLUSION

This study examines how agglomeration externalities affect regional wage disparities in Indonesia. The findings show that agglomeration significantly influences regional wages, confirming that spatial economic concentration remains an important source of wage inequality. Among localization economies, specialization positively affects wages, while diversity and competition have negative effects, consistent with MAR externality theory. Urbanization economies, measured by local density, also increase regional wages. These results suggest that regions with stronger economic concentration and specialization tend to offer higher wages. Worker characteristics also play an important role. Higher education, greater work experience, and stronger skills are associated with higher wages. Male and married workers also tend to earn higher wages, indicating persistent labor market segmentation and differences in household income patterns. Thus, regional economic structures and individual worker characteristics jointly shape wage disparities.

The findings have several policy implications. First, regional wage gaps can be reduced by promoting more balanced regional development, including investment in infrastructure and the development of secondary cities outside major metropolitan areas. Second, regional policies should encourage specialization based on local comparative advantages. Third, improving access to education and vocational training is essential to strengthen workforce skills and reduce wage gaps. Finally, policies promoting gender equality can support more inclusive regional development.

REFERENCES

- Akbar, F. (2022). Gender Wage Gap: Evidence from Employment in Informal Sector. *The Journal of Indonesia Sustainable Development Planning*, 3(2), 104–117. <https://doi.org/10.46456/jisdep.v3i2.301>.
- Barufi, A. M. B., Haddad, E. A., & Nijkamp, P. (2023). Urban Agglomeration, City Size, and Spatial Density Effects on Wage Flexibility: New Evidence on the Wage Curve in Brazil. *Regional Science Policy & Practice*, 15(9), 1998–2026. <https://doi.org/10.1111/rsp3.12669>.
- Beltrán Tapia, F. J., Díez-Minguela, A., & Martínez-Galarraga, J. (2018). Tracing the Evolution of Agglomeration Economies: Spain, 1860–1991. *The Journal of Economic History*, 78(1), 81–117. <https://doi.org/10.1017/S0022050718000086>.
- Capello, R. (2014). Proximity and Regional Innovation Processes: Is There Space for New Reflections? In Torre, A., & Wallet, F (Eds). *Regional Development and Proximity Relations*. Cheltenham: Edward Elgar Publishing.
- Catherine, S., Ebrahimian, M., Sraer, D., & Thesmar, D. (2022). Robustness Checks in Structural Analysis. *NBER Working Paper No. 30443*.
- Chauvin, J. P., Glaeser, E., Ma, Y., & Tobio, K. (2017). What is Different About Urbanization in Rich and Poor Countries? Cities in Brazil, China, India, and the

- United States. *Journal of Urban Economics*, 98, 17–49. <https://doi.org/10.1016/j.jue.2016.05.003>.
- Chen, A., Dai, T., & Partridge, M. D. (2021). Agglomeration and Firm Wage Inequality: Evidence from China. *Journal of Regional Science*, 61(2), 352–386. <https://doi.org/10.1111/jors.12512>.
- Chen, L., Hasan, R., & Jiang, Y. (2022). Urban Agglomeration and Firm Innovation: Evidence from Asia. *The World Bank Economic Review*, 36(2), 533–558. <https://doi.org/10.1093/wber/lhab022>.
- Chumaidiyah, E., Dewantoro, M. D. R., Fauzi, P. M., & Kamil, A. A. (2023). Selection of Industrial Sites Using a Web-Based Geographical Information System to Minimize Risks: A Case Study in West Java, Indonesia. *Sustainability*, 15(22), 16034. <https://doi.org/10.3390/su152216034>.
- Combes, P.-P., & Gobillon, L. (2015). The Empirics of Agglomeration Economies. In. Duranton, G., Henderson, J. V., & Strange, W. C. *Handbook of Regional and Urban Economics*, 247–348. Amsterdam: Elsevier.
- Elvidge, C. D., Zhizhin, M., Ghosh, T., Hsu, F.-C., & Taneja, J. (2021). Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019. *Remote Sensing*, 13(5), 922. <https://doi.org/10.3390/rs13050922>.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas. *International Journal of Climatology*, 37(12), 4302–4315. <https://doi.org/10.1002/joc.5086>.
- Groot, S., de Groot, H. L. F., & Smit, M. J. (2014). Regional Wage Differences in The Netherlands: Micro Evidence on Agglomeration Externalities. *Journal of Regional Science*, 54(3), 503–523. <https://doi.org/10.1111/jors.12070>
- Grover, A., Lall, S., & Timmis, J. (2023). Agglomeration Economies in Developing Countries: A Meta-Analysis. *Regional Science and Urban Economics*, 101, 103901. <https://doi.org/10.1016/j.regsciurbeco.2023.103901>.
- Hu, S., Xiang, W., & Wan, Y. (2026). Digital Financial Inclusion and Economic Growth: Multi-Dimensional Evidence from Coverage, Depth, and Digitisation. *Journal of Risk and Financial Management*, 19(4), 284. <https://doi.org/10.3390/jrfm19040284>.
- Ikeda, S. (2020). The Economy of Cities: Jane Jacobs's Overlooked Economic Classic. *The Independent Review*, 24(4), 605–618.
- Kleibergen, F., & Paap, R. (2006). Generalized Reduced Rank Tests Using the Singular Value Decomposition. *Journal of Econometrics*, 133(1), 97–126. <https://doi.org/10.1016/j.jeconom.2005.02.011>.
- Maghfirah, M., & Samosir, O. B. (2022). Wages of Workers Spatial Analysis in Indonesia Region 2019. *Proceedings of The International Conference on Data Science and Official Statistics*, 2021(1), 708–716. <https://doi.org/10.34123/icdsos.v2021i1.66>.

- Özgüzel, C. (2023). Agglomeration Effects in a Developing Economy: Evidence from Turkey. *Journal of Economic Geography*, 23(4), 823–846. <https://doi.org/10.1093/jeg/lbac035>.
- Prasertsoong, N., & Puttanapong, N. (2022). Regional Wage Differences and Agglomeration Externalities: Micro Evidence from Thai Manufacturing Workers. *Economies*, 10(12), 319. <https://doi.org/10.3390/economies10120319>.
- Putra, A. A., Hasibuan, H. S., Tambunan, R. P., & Lautetu, L. M. (2024). Integration of the Sustainable Development Goals into a Regional Development Plan in Indonesia. *Sustainability*, 16(23), 10235. <https://doi.org/10.3390/su162310235>.
- Ranta, M., & Ylinen, M. (2024). Employee Benefits And Company Performance: Evidence from a High-Dimensional Machine Learning Model. *Management Accounting Research*, 64, 100876. <https://doi.org/10.1016/j.mar.2023.100876>.
- Ridhwan, M. M. (2021). Spatial Wage Differentials and Agglomeration Externalities: Evidence from Indonesian Microdata. *Economic Analysis and Policy*, 71, 573–591. <https://doi.org/10.1016/j.eap.2021.06.013>.
- Rismaya, E., & Yusuf, A. A. (2026). Persistence of Labor Market Disparities in Indonesia: An Analysis of Regional Productivity and Wage Gaps, 2018–2023. *Journal of Society and Development*, 6(1), 14–22. <https://doi.org/10.57032/jsd.v6i1.339>.
- Rotman, A., & Mandel, H. (2023). Gender-Specific Wage Structure and the Gender Wage Gap in the U.S. Labor Market. *Social Indicators Research*, 165(2), 585–606. <https://doi.org/10.1007/s11205-022-03030-4>.
- Sangkasem, K., & Puttanapong, N. (2022). Analysis of Spatial Inequality Using DMSP-OLS Nighttime-Light Satellite Imageries: A Case Study of Thailand. *Regional Science Policy & Practice*, 14(4), 828–849. <https://doi.org/10.1111/rsp3.12386>.
- Small, D. S., Tan, Z., Ramsahai, R. R., Lorch, S. A., & Brookhart, M. A. (2017). Instrumental Variable Estimation with a Stochastic Monotonicity Assumption. *Statistical Science*, 32(4), 561–579. <https://doi.org/10.1214/17-STS623>.
- Sun, H. (2026). The Impact of Vocational Skills Training on Earnings: Evidence from China. *International Review of Economics & Finance*, 107, 105079. <https://doi.org/10.1016/j.iref.2026.105079>.
- Tabash, M. I., Elsantil, Y., Hamadi, A., & Drachal, K. (2024). Globalization and Income Inequality in Developing Economies: A Comprehensive Analysis. *Economies*, 12(1), 23. <https://doi.org/10.3390/economies12010023>.
- Tan, M., Li, X., Li, S., Xin, L., Wang, X., Li, Q., Li, W., Li, Y., & Xiang, W. (2018). Modeling Population Density based on Nighttime Light Images and Land Use Data in China. *Applied Geography*, 90, 239–247. <https://doi.org/10.1016/j.apgeog.2017.12.012>.
- Umair, M., Ahmad, W., Hussain, B., Antohi, V. M., Fortea, C., & Zlati, M. L. (2024). The Role of Labor Force, Physical Capital, and Energy Consumption in Shaping

Agricultural and Industrial Output in Pakistan. *Sustainability*, 16(17), 7425. <https://doi.org/10.3390/su16177425>.

Yang, Z., Hong, Y., Guo, Q., Yu, X., & Zhao, M. (2022). The Impact of Topographic Relief on Population and Economy in the Southern Anhui Mountainous Area, China. *Sustainability*, 14(21), 14332. <https://doi.org/10.3390/su142114332>.