Signifikan: Jurnal Ilmu Ekonomi
Volume 12 (1), 2023: 27 - 44
P-ISSN: 2087-2046; E-ISSN: 2476-9223

Urban Size and Labor Market Premium: Evidence from Indonesia

Msy. Nourma Yunita Sari[1*], Muhammad Halley Yudhistira[2]
[1]BPS - Statistics of Tangerang Municipality, Indonesia
[2]Department of Economics, Universitas Indonesia, Indonesia
E-mail: [1]msy.nourma@gmail.com, [2]m.halley@ui.ac.id
[*]**Corresponding author**

JEL Classification:
J24
J31
O18

Received: 09 September 2022

Revised: 12 January 2023

Accepted: 14 January 2023

Available online: 15 April 2023

Published regularly: April 2023

Abstract

The study of economic agglomeration is again a concern in the urban economic literature, especially in describing urban areas and better econometric approaches. This study improves the size of cities to become urban and suburban, reflecting the flow of commuting, using the 2010 and 2015 Landscan data to measure economic density better and reduce bias due to measurement errors. Empirically, using this density and using the 2SLS estimation technique with instrument variables in the form of earthquake risk and ruggedness measures, the result of a city twice as large can increase wages by 61 percent. This result is higher than most other literature because the sample only covers urban areas. This study also shows that workers with characteristics such as higher education, the formal sector, the service sector, and white-collar jobs get more enormous benefits in urban areas.

Keywords:
agglomeration economies; economic density; urbanization

How to Cite:
Sari, M. N. Y., & Yudhistira, M. H. (2023). Urban Size and Labor Market Premium: Evidence from Indonesia. *Signifikan: Jurnal Ilmu Ekonomi*, *12*(1), 27-44. https://doi.org/10.15408/sjie.v12i1.27999.

INTRODUCTION

In 2018, more than half of the world's population lives in cities, 55.27 percent of the total 4.19 billion-world population. This figure will continue to increase, estimated to reach 68 percent in 2050 (United Nations, 2018). The benefits of living in cities may partly drive this imbalance in the concentration of rural and urban populations. For workers, higher wages and urban amenities can beat all the congestion costs that arise in cities, such as high prices for land, housing, environmental problems, especially pollution, transportation costs, and longer commute times. Likewise, for companies, higher wages may cause the cost of production input to be more expensive. However, these companies still choose to operate in urban areas, not moving to lower fees.

The foremost contribution of this study is in being one of the first in the literature to causally estimate the impact of urban size on the individual's labor market outcomes in developing countries by determining the urban size with a clustering algorithm so that it can reflect economic density. Economic density classified as urban and suburban, to reduce measurement error. In doing so, this paper adds to an existing body of work on benefit from the urban size on the individual's labor market outcomes. Outside of this work, much of the existing research on urban labor market premium has focused on developed countries such as United States (Glaeser & Mare, 2001; Glaeser & Resseger, 2010; Baum-Snow & Pavan, 2012; Fu & Ross, 2007; Yankow, 2006), France (Combes, 2008; Combes et al., 2012), United Kingdom (D'Costa & Overman, 2014), Spain (Puga, 2017), Europe (Ciccone, 2002), and several other works of literature and urban size data based on administrative boundaries, which cannot yet describe economic density and is currently considered inappropriate in determining city size. As for developing countries, the study is still limited. The constraints are the limitation of longitudinal data and urban size data, which cannot yet describe economic density.

The labor market in developing countries is quite different from developed countries. Large informal sector and labor with low skills characterize this labor market. The benefit of a higher population and business concentration in urban areas is higher for formal industry and highly skilled workers. Therefore, the impact of urban size on labor market outcomes in developing countries may differ from that of developed countries, so it is interesting to study. Recently, there has been more work on developing countries with, for example, papers on Columbia (Duranton, 2016), China, Brazil, and India (Chauvin, 2017), and Sub-Saharan Africa (Henderson, 2019).

In developed countries, urban size reflects the economic density determined by commuting flows (Duranton, 2015). However, this cannot be used in developing countries because of the limitations of the data. City size based on administrative boundaries has not been able to capture de facto economic density (Widya et al., 2019; Khairunurrofik, 2017). Due to pre-determined administrative units, cities, or villages, based on qualitative aspects of land use, which can include areas that don't have populations or rural areas and cities usually grow beyond their administrative boundaries. However, this cannot be done in developing countries due to limited data.

http://journal.uinjkt.ac.id/index.php/signifikan
https://doi.org/10.15408/sjie.v12i1.27999

This study evaluates the impact of city size on the individual labor market premium using the Indonesian case, due to 3 reasons. First, Indonesia is one of the developing countries with the characteristics of a labor market dominated by the informal sector and low-skilled labor. Second, Indonesia is the fourth most populous country in the world, experiencing a relatively fast rate of urbanization. Third, Indonesia will also experience a window of opportunity, with an enormous population growth of productive age. It is hoped that this study can provide useful results for policymakers to face this big opportunity. This study also conducts a more in-depth heterogeneity analysis of several characteristics of workers because the benefits of city size are not the same for every worker (Duranton, 2015). Estimating the impact of city size on labor market outcomes is quite challenging due to endogeneity problems that cause the estimation results using the Ordinary Least Square (OLS) method to be biased. Endogeneity problems at the city level occur due to the non-random process of city size, in the form of population density, because of the unobserved attributes of the city, both in the form of natural amenities and urban public infrastructure encourages urban development and also boosts productivity. Besides, the problem of reverse causality is that the average wage in urban areas increases the number of workers, and density increases, and the measurement error of urban size. Therefore, to reduce this bias, we use the Two-Stage Least Square (TSLS) method. Besides, this study improve the city's size to become urban and suburban that reflects the commuting flow so that they can better measure economic density and reduce bias due to measurement error. At the same time, the endogeneity problem at the individual level occurs due to select sorting driven by unobserved individual characteristics. Individuals with high abilities can live in larger cities and, at the same time, affect the wages received. So this study controls education and age as proxies of experience, other observed individual characteristics, job control, and the fixed effect industry are carried out to factor abilities by conditioning education.

At present, the study of economic agglomeration is again a concern in the urban economic literature, especially in describing urban areas and better econometric approaches, resulting in a more precise estimate of the impact of city size on the labor market premium. It is because advancing data has made it possible to do this. Data on a better spatial scale makes it possible to obtain an urban size that matches the commuting flow. By utilizing geospatial data that is increasingly developing today, several economists and geographers have conducted studies on urban sizes that are in line with commuting's flow. As far as the researchers know, there are several approaches proposed, including agglomeration index (Uchida & Nelson, 2009), urban clustering (Henderson, 2019), night-time lights (Dingel, 2019), and finally using building information with an algorithm machine learning (Arribas-Bel & Lopez, 2019; Bellefon et al., 2019). This study uses an urban cluster approach such as that conducted by Henderson, 2019, for Sub-Saharan Africa. Using Landscan 2012 data to determine the threshold that matches the commuting flow, Henderson defines the city and town. Henderson concluded that this measure was perfect in estimating economic density measures.

This research contribution includes two things. First, determine the city's size with a clustering algorithm, based on the population density and total population threshold consistent with commuting flow so that it can reflect economic density, classified as urban and suburban, to reduce measurement error. Second, this study examines causal impact of city size on the individual labor market premium.

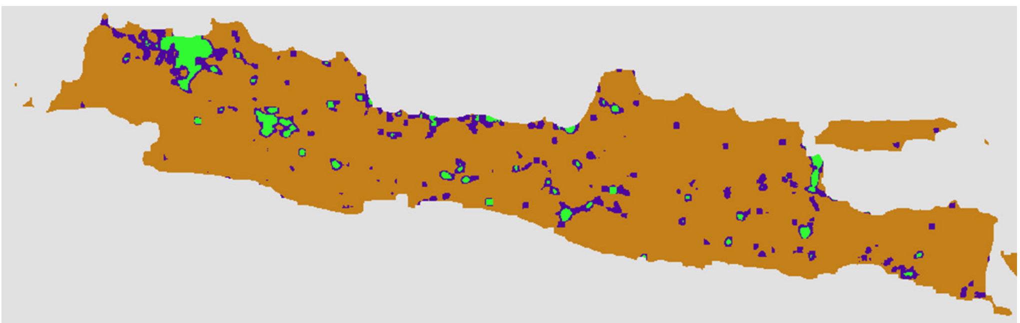
METHODS

Data

Wage and worker's characteristics - Our main data source for for worker's wage and individual characteristics is 2010 and 2015 National Labor Force Survey (Sakernas) from Central Bureau of Statistics. The sample is restricted to working-age residents of 15-64 years, with wages not equal to zero. This research is limited to the net income of workers with self-employed status, casual workers in agriculture, casual workers in non-agriculture, and net wages, either in cash or in-kind, for workers with labor/employee/employee status. Workers working in the agriculture, forestry sectors, and mining and quarrying sectors were excluded from the sample.

Urban density - We construct urban density from 2010 and 2015 Landscan data sourced from Oak Ridge National Laboratory. As the urban definition is likely to be beyond the administrative boundaries, we follow Djikstra & Poelman's method (2014) in harmonizing the definition of urban and rural areas. They applied a cluster algorithm using population grid data with a resolution of 1 km^2. Adjacent population grids are categorized as urban areas if each population grid meets a certain population density threshold and population threshold. We follow Henderson's strategy (2019) in determining the status of urban areas using different threshold. A settlement area may fall into one of four categories, i.e. (a) Core, with a population density of more than 4000 people per km^2 and a total population of more than 400,000 people, (b) fringe, with a population density of more than 2000 people per km^2, (c) suburban, a standing area alone, a population density of more than 2000 people per km^2 and a total population of more than 200,000 people, and rural if otherwise. An urban area is then defined as those who have cores and fringes. Figure 1 illustrates the urban areas in Java Island, the most populous island in Indonesia.

Figure 1. Urbanized areas in Java Island, Indonesia



Note: Green area represents core, an area with a population density of more than 4000 people per km^2 and a total population of more than 400.000 people. Purple area represents suburban area, an area with a population density of more than 2000 people per km^2 and a total population 200.000 people. Brown area represents rural area, an area with a population density of less than 2000 people per km^2 or total population of less than 200.000 people

http://journal.uinjkt.ac.id/index.php/signifikan
https://doi.org/10.15408/sjie.v12i1.27999

Instrumental variables - To construct the instrument variables, we rely on district/city-level earthquake risk index from the Indonesian Disaster Risk Index by the Indonesian Disaster Management Agency. The index is published in 2014 and represents information of the level of vulnerability of the area at the territory of the State of Indonesia. Disaster Risk Index compiled based on hazard components, loss, and capacity. Component hazards arranged by intensity and probability parameters incident. The loss components are compiled based on socio-cultural parameters, economic, physical, and environmental. Capacity components are arranged based on capacity parameters regulations, institutions, systems early warning, education, training, skills, mitigation, and systems preparedness. At last ruggedness measure is constructed from Shuttle Radar Topography Mission by the United States Geological Survey. We follow Riley (1999) to construct the index.

Additional variables – We also rely on other source of data for our robustness test on model specification. We rely on the 2010 Population Census (SP) and the 2015 Intercensal Population Survey (SUPAS) for constructing the population density at district level.

Identification Strategy

The empirical model in this study uses individual data (i) in the city $c(i)$ in the year (t) period. Mathematically, the empirical model can be written with

$$y_{ic(i)t} = \alpha_0 + \alpha_1 \log Density_{c(i)t} + \alpha_2 X_{ic(i)t} + \alpha_3 A_{c(i)t} + \delta_t + \varepsilon_{ic(i)t} \qquad (1)$$

The dependent variable $(y_{ic(i)t})$ represents a set of urban individual labor market outcomes, including employment opportunities in urban areas, the wage premium for those working, and the absorption of formal workers in urban areas. The wage used is the amount nominal wages or income per month received in the form of money and goods. Differences in nominal wages reflect differences in productivity more than differences in real wages that better reflect differences in living standards. The variable of interest is urban size. The measure used is population density, the total population divided by area, both obtained from 2010 and 2015 Landscan data using cluster algorithms or with administrative boundary approaches. It is important to note that $c(i)$ shows the city where the individual works, not the city of residence.

The variable of interest is urban size. The measure used is population density, the total population divided by area, both obtained from 2010 and 2015 Landscan data using cluster algorithms or with administrative boundary approaches. The area used is also derived from Landscan data processing. The use of these two variables in equation (3.1) reduces bias because the area is often not well measured (Duranton, 2020). According to Duranton (2015), the use of density measures is more robust against zoning to provide more reliable results. It is important to note that $c(i)$ shows the city where the individual works, not the city of residence.

X_{it} is a vector of individual characteristics in year t, including age, age squared, duration of work in current jobs, and it's square as a rough proxy of experience. Variables reflect skills, namely years of schooling. Besides, other individual characteristic variables used are dummy gender, hours of work and its square, employment field (Classification of Standard Indonesian Business Field 2005 (KBLI 2005) 2 digits, and position classified according to Indonesian Classification Classification (KJI 2002) 1 digits. $A_{c(i)t}$ is a vector of city characteristics in year t, including human capital city and island dummy. Clustered standard clustered errors based on KBLI are used to accommodate the correlation between individuals in the same KBLI.

An empirical strategy in estimating the effect of urban size on individual labor market premiums needs to consider the endogeneity problem, both at the individual level and at urban area level. The endogeneity problem at the individual level occurs because unobserved individual characteristics drive individual sorting. Individuals with high ability tend to choose to live in a larger city, so this study controls education and age as proxies of experience, other observed individual characteristics, and the occupation and industry fixed effect. The use of occupation and industry fixed effect as a way of factoring abilities by conditioning occupation because specific jobs or sectors correspond to particular skills. Capturing work done by workers, which is also an effect of past careers, can be considered a more correlated measure with current skills than education (Combes, 2015). Besides, job control and industry fixed effects can eliminate the benefits of a wider choice of jobs and industries when moving to a larger city (Henderson, 2019).

The issue of endogeneity at the urban area level occurs due to the non-random process of city size, in the form of population density, due to the city's unobserved attributes, both in the form of natural amenities and city public infrastructure that encourage city development and also boost productivity. Besides, there is also the problem of reverse causality, where a larger average wage in urban areas attracts workers to the city so that the density increases. To reduce this bias, the Two-Stage Least Square (TSLS) method. Instrument variables (IV) used in this study are the interaction between district/ city earthquake risk index and the district/city ruggedness. Earthquake risk and ruggedness measures are exogenous events, which are important determinants of population settlement patterns (Combes, 2010; Duranton & Turner, 2018), and both variables do not directly determine the individual's wage level. When the terrain is rough, then construction on an inclined side is more expensive than a flat area, so the rough terrain naturally encourages diffuse development. It can also mean a spread population. In contrast, flat areas or high mountains tend to make development more densely populated to become denser (Burchfield, 2006). As for earthquake risk, areas with a higher earthquake risk have a lower population density than areas with low earthquake risk. Our first-stage equation is then as follows:

$$\log Density_{c(i)t} = \beta_0 + \beta_1 Earthquake_{c(i)t} * ruggedness_{c(i)t} + \beta_2 X_{ic(i)t} + \beta_3 A_{c(i)t} + \mu_t + u_{ic(i)t} \quad (2)$$

Where $earthquake_{c(i)t}$ and $ruggedness_{c(i)t}$ are earthquake risk score and ruggedness index, respectively.

http://journal.uinjkt.ac.id/index.php/signifikan
https://doi.org/10.15408/sjie.v12i1.27999

RESULT AND DISCUSSION

First stage: Geological characteristics and population density

Table 1 provides the first-stage result of the earthquake risk score and ruggedness index on the population density. A one-point increase in an interaction between earthquake risk score and ruggedness index corresponds to 2.9 percent decrease in the district's population density.

Table 1. F*irst Stage Result*: Earthquake Risk Index and Ruggedness on Population Density

Variable	(1)	(2)	(3)	(4)
Earthquake score x ruggedness index	**-0.0283 (0.0023)**	**-0.0287 (0.0023)**	**-0,0286*** (0.0024)**	**-0.0290*** (0.0024)**
Individual characteristics	N	N	Y	Y
Urban characteristics	N	Y	N	Y
Province-year fixed effect	Y	Y	Y	Y
Occupation and industry fixed effect	Y	Y	Y	Y
Mean of population density				
R-sq	0.326	0.330	0.326	0.329
Observations	108,831	108,831	103,984	103,984

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively. Standard errors in parentheses are clustered at industry classification level. Samples are limited to those working in urban and suburban area. Individual characteristics includes age, age squared, years of schooling, experience, experience squared, working hours, working hours squared, and gender dummy. Urban characteristics include average human capital.

Second stage: Effects of urban density on individual wages

We now move to our main result. Table 2 shows the OLS and TSLS estimation results from equation (1). The results obtained are that urban size has a positive and significant impact at the 1 percent level, both on the OLS and 2SLS estimates. The bigger the size of a city, the bigger the wages will be. The city's size in this study is population density because city density is a robust measure of zoning. However, the use of the total population size will still be carried out in the robustness test.

Due to endogeneity problems, this study uses the TSLS method. The OLS estimation results in column 1 and column 2 are only for descriptive purposes. In Table 2, there are the results of the first stage F statistics for the instrument. The results passed the test for weak instruments Stock and Yogo (2005) shown from the first stage-F-statistic value exceeding the relative bias critical value measure Stock-Yogo. To overcome the endogeneity problem due to individual sorting, which is driven by unobserved individual characteristics, the study controls variables of education (school duration) and experience (age and length of work in current work). Also, control occupation and industry fixed effect because a particular job or sector corresponds to a particular ability. Also, occupation and industry fixed effects can eliminate the benefits of a more extensive choice of occupations and industries when moving to a larger city (Henderson, 2019).

https://doi.org/10.15408/sjie.v12i1.27999

The 2SLS estimation results using control occupations and industries fixed effect (column 4) show that an increase in population density by 1 percent increases urban wages by 61 percent. Compared to column (2), the OLS estimate underestimates because of the omitted bias variable. Compared with column (3), the estimation results without controlling the occupation and the industry fixed effect are overestimated because of the omitted bias variable (ability).

Table 2. Estimation result: Impact of Urban Size on Wages

Variable dependent: : ln (wage)	(1)	(2)	(3)	(4)
	OLS	OLS	2SLS	2SLS
Ln (population density)	**0.041*** (0.019)**	**0.040*** (0.019)**	**0.830*** (0.136)**	**0.615*** (0.104)**
Individual characteristics	Y	Y	Y	Y
Urban characteristics	Y	Y	Y	Y
Province-year fixed effect	Y	Y	Y	Y
Occupation and industry fixed effect	N	Y	N	Y
Mean of population density	6.741,007	6.741,007	6.741,007	6.741,007
First stage F-stat	-	-	136.9	147.2
Observations	103,984	103,984	103,984	103,984

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively. Standard errors in parentheses are clustered at industry classification level. Samples are limited to those working in urban and suburban area. Individual characteristics includes age, age squared, years of schooling, experience, experience squared, working hours, working hours squared, and gender dummy. Urban characteristics include average human capital. 2SLS estimation method uses the interaction between earthquake risk score and ruggedness index as the instrument.

This result is much higher than the results obtained from the economic agglomeration literature for developed countries such as Combes (2010) for France of 0.027, Puga (2017) for Spain, the resulting short-term urban wage elasticity of 0.024 and 0.051 for the medium term, Glaeser (2001) for the United States, using longitudinal data, the wage premium is 0.039. This result is also higher than other developing countries, such as Duranton (2016) for Columbia; the results obtained are 0.054. However, Duranton does include urban areas and rural areas so that these results cannot be compared directly. Chauvin (2017) found wage premium results for China of 0.323 and India of 0.208. Meanwhile, for Indonesia, Bosker (2019) found that the urban wage premium produced varies depending on the approach to defining the metropolitan area used. With the cluster algorithm approach, the resulting wage impact is 0.49, while the nights time light approach results in an impact of 0.796. Henderson (2019) for Sub-Saharan Africa, with a similar approach to the one this study uses to define urban areas, finds results for urban wage premium of 0.169 and urban household income premium of 0.523. According to Henderson, the magnitude of the estimation results obtained is due to the size of the city used, namely density is the right size produced based on a particular area, in contrast to the area based on administrative boundaries, which is very likely to

http://journal.uinjkt.ac.id/index.php/signifikan
https://doi.org/10.15408/sjie.v12i1.27999

cause measurement errors. Henderson proves this by entering the total population size together with the area because the area is correlated with the total population so that there is no omitted bias variable. The result is that the resulting coefficient is very close in absolute value, only has a contradicting sign. Besides, this higher result may indicate that there is still some bias because the sorting problem in the time-invariant unobserved individual characteristics has not been appropriately resolved.

An interesting policy related to the current density of urban areas in Indonesia is the plan to relocate the Indonesian capital, which is currently in DKI Jakarta Province, to East Kalimantan Province. If the plan is carried out, the population that usually does activities in Jakarta will certainly decrease, for example, by around 10 percent. Reducing the density of DKI Jakarta by 10 percent certainly changes not only the economic activities that occur in Jakarta but also its productivity, which is shown by wages, because the benefits of externalities created will certainly decrease. Based on the main results of this study, the wage elasticity of population density is 0.615, so a 10 percent reduction in Jakarta's density has the potential to eliminate Jakarta's wage premium by 6.1 percent. Meanwhile, East Kalimantan Province, especially Kutai Kartanegara Regency and North Penajam Paser Regency, as the location centers for the new capital plan, may not necessarily be able to create premium benefits that are comparable to the potential loss of premium that has occurred in Jakarta.

Heterogeneity Analysis

Heterogeneity analysis needs to be done because city size benefits are not the same for every worker (Duranton, 2015). Equation (1) estimates the average effect of the urban density on the individual's wage. This section will evaluate the potential for heterogeneity between sample groups based on gender, age group, education, formal-informal jobs, type of work, and work sector. First, differentiation between men and women is carried out because the labor market for men and women is often different. The gap between gender wages occurs in almost all countries in the world, including Indonesia. Second, differentiation by age group is due to the possibility of differences in urban benefits received by different age groups. There is quite a lot of literature that proves that wages grow with the city's length of life, or that there are learning benefits that workers receive overtime. This theory predicts more significant benefits for workers who stay longer, arguing that older workers receive more. Therefore, the learning mechanism can be proven indirectly by differentiating these age groups (Duranton, 2005). Third, differentiation based on education, the theory predicts cities to be more profitable for individuals who have high education and skills (Duranton, 2015). Fourth, the distinction is also made based on formal-informal jobs. The theory predicts that formal sector workers, especially those with high skills, feel more significant urban benefits. Meanwhile, informal workers dominate labor market conditions in Indonesia. Fifth, differentiation based on type of work, white-collar that reflects professional jobs with high skills and blue collar that reflects manual workers with low skills. Sixth, the impact of city size can also be heterogeneous between industries depending on the industry's characteristics. In a meta-analysis, Melo (2009) concluded that city size's average impact tends to be stronger for the manufacturing sector than for the services sector.

From Table 3, Panel A, differentiation based on sex, the estimation results show that men get wage premium working in urban areas, which is much higher than women. These results are found in the literature on wage disparity studies between genders. A study conducted by Nordman et al. (2011) for West Africa found that wage disparities that occur are due to education gaps and differences in sector allocations, which can explain one-third of this gender gap. This result is also consistent with a study conducted by Sohn (2015) for Indonesia, where women earn 30 percent less than men, both for paid workers and self-employed workers. The gender gap explained by individual and occupational characteristics is only a quarter gap for paid workers and half the gap for self-employed workers. Although the sizeable unexplained gap cannot be discriminatory, the researchers speculate that this gap is driven by culture, religion, and social norms in Indonesia that restrict women from remaining in the domestic sphere.

Table 3. Estimation Result: Sample groups

	Ln (population density)		Obs	First stage F- stat
	Coef	Std. error		
Panel A: Gender				
Men	0,858 ***	0,101	66124	118,1
Women	0,302***	0,121	37860	120,2
Panel B: Age group				
Age 20-30	0,598***	0,145	27081	109
Age 40-50	0,617***	0,106	29148	202
Panel C : Education				
Less than high school	0,497***	0,099	87365	118,8
More than high school	0,907***	0,209	16619	125,1
Panel D : Formal Informal				
Formal	0,717***	0,139	70322	91,6
Informal	0,383***	0,102	33662	136,7
Panel E : Type of work				
White_collar	0,890***	0,236	25890	107
Blue_collar	0,378***	0,125	53044	60,5
Panel F : Work sector				
Industrial	0,476***	0,173	33512	24,8
Services	0,756***	0,116	70472	829

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively. Standard errors in parentheses are clustered at industry classification level. Samples are limited to those working in urban and suburban area. Individual characteristics includes age, age squared, years of schooling, experience, experience squared, working hours, working hours squared, and gender dummy. Urban characteristics include average human capital. 2SLS estimation method uses the interaction between earthquake risk score and ruggedness index as the instrument.

http://journal.uinjkt.ac.id/index.php/signifikan
https://doi.org/10.15408/sjie.v12i1.27999

In panel C, the distinction is made based on individual education. The results show that individuals with higher education benefit from a more significant urban wage premium. These results are consistent and support the theory that economic agglomeration has more impact on workers with high education and skills (Glaeser & Resseger, 2010; Bacolod, 2009). Different results were found by Duranton (2005) for Columbia, a smaller wage premium for highly educated workers in large cities due to the informal sector's high premium in urban areas, which is dominated by young workers and with low education.

In panel D, differentiation is based on employment in the formal or informal sector. The estimation results show that the benefits of urban wage premium are more significant for workers in the formal sector. However, workers in the informal sector also benefit from a large urban wage premium. This result is consistent with the results found in developed countries, but contradicts the results found by Duranton (2005).

In panel E, the distinction is made based on the type of work, and the results show that a white-collar that reflects professional work gets the most significant wage premium. Workers with high education and skills dominate professional workers. These workers benefit significantly in the form of learning from interactions with their fellow environment. In contrast to white collars, blue collars, which are synonymous with manual labor, benefit from a lower wage, such as in informal work, this type of work also benefits from a large urban wage premium. Bacolod et al. (2009) concluded that cognitive skills were more valuable in cities, while motor skills and physical strength were less valued. Gould (2007) also found that the benefit of city size exists for white collars but not for blue collars.

In panel F, the differentiation is made based on the industrial sector and the service sector. The service sector has more significant urban benefits than the industrial sector, contrary to the conclusion of Melo (2009). It may be due to the city's increasing size, which reflects the diversity in the city. The benefits of industrial diversity in this city are more profitable for the service sector than in the industrial sector. The activities carried out by this sector are closely related to the locations of other sectors, such as in the trade, hotels and restaurants sector, and the transportation and communication services sector. This result is consistent with the results found by Brulhart & Mathys (2008) for Europe, and the financial services sector gets more urban benefits than the manufacturing sector, this also occurs in Indonesia when disaggregated into smaller subsectors.

CONCLUSION

This study estimates the magnitude of city size's impact on wages by evaluating different approaches to defining urban areas. This study selects Indonesia, a developing country with a large population and a fast rate of urbanization. By selecting the right threshold, the resulting urban area is quite good at capturing economic density, reflecting commuting flows. Besides, the city's size is based on population density obtained from the 2010 and 2015 Landscan data. This density measure is more precise than the population density obtained based on administrative boundaries because the area generated from the

Landscan data is more accurate. The results obtained from estimating the impact of city size on wage premium give an extraordinary effect, higher than most of the rest of the literature, because the sample only includes urban areas and uses a more precise measure of density than based on administrative boundaries.

Hence, the problem of measurement error is very likely to be adequately resolved. Besides, this higher result may indicate that there is still some bias because the sorting issue in the time-invariant unobserved individual characteristics has not been appropriately resolved. Heterogeneity analysis shows that the benefits of urban wage premiums received by men are more significant than those received by women, *workers with characteristics such as higher education, the formal sector, the service sector, and white collar jobs* get more enormous benefits in urban areas. This result also suggests that cognitive skills are more valuable in cities, while motor skills and physical strength are less valued. The policy implications of this research are to increase the cognitive skills of workers and strengthen the service sector.

REFERENCES

Ahlfeldt, G., & Pietrostefani, E. (2019). The Economic Effects of Density. *Journal of Urban Economics*, *111,* 93-107. https://doi.org/10.1016/j.jue.2019.04.006.

Arribas-Bel, D., & Garcia-Lopez, M.-A. (2019). Building(s and) Cities: Delineating Urban Areas with A Machine Learning Algorithm. *Journal of Urban Economics*, *125,* 103217. https://doi.org/10.1016/j.jue.2019.103217.

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

Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Sorting, Selection, and Agglomeration. *Journal of Political Economy*, *122*(3), 507-553. https://doi.org/1086/675534.

Bellefon, M.-P. d., Combes, P.-P., Duranton, G., Gobillon, L., & Gorin, C. (2019). Delineating Urban Areas Using Building Density. *Journal of Urban Economic, 125,* 103226.https://doi.org/10.1016/j.jue.2019.103226.

Bosker, M., Park, J., & Roberts, M. (2019). Definition Matters. Metropolitan Areas and Agglomeration Economies in A Large Developing Country. *Journal of Urban Economics, 125,* 103275. https://doi.org/10.1016/j.jue.2020.103275.

Brülhart, M., & Mathys, N. (2008). Sectoral Agglomeration Economies in a Panel of European Regions. *Regional Science and Urban Economics*, *38*(4), 348-362. https://doi.org/10.1016/j.regsiurbeco.2008.03.003.

Burchfield, M., Overman, H., Puga, D., & Turner, M. (2006). Causes of Sprawl: A Portrait From Space. *The Quarterly Journal of Economics*, *121*(2), 587-633. https://doi.org/10.1162/qjec.2006.121.2.587.

Chauvin , J., 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.

http://journal.uinjkt.ac.id/index.php/signifikan
https://doi.org/10.15408/sjie.v12i1.27999

Ciccone, A. (2002). Agglomeration Effects In Europe. *European Economic Review*, *46*(2), 213-227. https://doi.org/10.1016/S0014-2921(00)00099-4.

Combes, P.-P., & Gobilon, L. (2015). The Empirics of Agglomeration Economies. *Handbook of Regional and Urban Economics*, *5,* 247-348. https://doi.org/10.1016/B978-0-444-59517-1.00005-2.

Combes, P.-P., 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, P.-P., Duranton, G., Gobilon, L., & Roux, S. (2010). Estimating Agglomeration Economies with History, Geology, and Worker Effects. *Working Paper The University of Chicago Press*, 15-66.

D'Costa, S., & Overman, H. (2014). The Urban Wage Growth Premium: Sorting or Learning? *Regional Science and Urban Economics*, *48,* 168-179. https://doi.org/10.1016/ j.regsiurbeco.2014.06.006.

Dijkstra, L., & Poelman, H. (2014). A Harmonised Definition of Cities and Rural Areas: the New Degree of Urbanisation. *Working papers, Directorate-General for Regional And Urban Policy*.

Dingel, J., Miscio, A., & Davis, D. (2019). Cities, Lights, and Skills in Developing Economies. *Journal of Urban Economies*, *125,* 103174. https://doi.org/10.1016/j.jue. 2019.05.005.

Duranton, G. (2015). Growing through Cities in Developing Countries. *The World Bank Research Observer*, *30*(1), 39-73.

Duranton, G. (2016). Agglomeration Effects in Colombia. *Journal of Regional Science*, *56*(2), 210-238. https://doi.org/10.1111/jors.12239.

Duranton, G., & Puga, D. (2003). Micro-foundation of Urban Agglomeration Economies. *Handbook of Regional and Urban Economics, 4,* 2063-2117. https://doi.org/10.1016/S1574-0080(04)80005-1.

Duranton, G., & Puga, D. (2020). The Economics of Urban Density. *Journal of Economic Perspectives*, *34*(3), 3-26.

Duranton, G., & Turner, M. (2018). Urban Form and Driving: Evidence from US Cities. *Journal of Urban Economics*, *108,* 170-191. https://doi.org/10.1016/j.jue.2018.10.003.

Fu, S., & Ross, S. L. (2004). Wage Premia in Employment Clusters: Agglomeration Economies or Worker Heterogeneity. *Journal of Labor Economics, 31*(2), 271-304. https://doi.org/10.1086/668615.

Giuliano, G. (2019). Agglomeration Economies and Evolving Urban Form. *The Annals of Regional Science*, *63,* 377-398. https://doi.org/10.1007/s00168-019-00957-4.

Glaeser, E., & Gottlieb, J. (2009). The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium. *Journal of Economic Literature*, *47*(4), 983-1028.

Glaeser, E., & Mare, D. (2001). Cities and Skills. *Journal of Labor Economics*, *19*(2), 316-342. https://doi.org/10.1086/319563.

Glaeser, E., & Resseger, M. (2010). The Complementary Between Cities and Skills. *Journal of Regional Science*, *50*(1), 221-244. https://doi.org/10.1111/j.1467-9787.2009.00635.x.

Gould, E. D. (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.

Henderson, J., Nigmatulina, D., & Ktiticos, S. (2019). Measuring Urban Economic Density. *Journal of Urban Economics, 125,* 103188. https://doi.org/10.1016/j.jue.2019.103188.

Jones, Gavin; Rangkuti, Hasnani; Utomo, Ariane; McDonald, Peter;. (2016). Migration, Ethnicity, and Education Gradient in the Jakarta Mega Urban Region: A Spatial Analysis. *Bulletin of Indonesian Economic Studies*, *52*(1), 55-76. https://doi.org/10.1080/ 00074918.2015.1129050.

Khoirunurrofik. (2018). Local Economic Structure, Productivity Growth, and Industry Life Cycle: Evidence from Indonesia. *Asia-Pacific Journal of Regional Science, 2,* 453-475. https://doi.org/10.1007/s41685-018-0072-6.

Lamorgese, A. R., Olivieri, E., & Paccagnella, M. (2019). The Wage Premium in Italian Cities. *Italian Economic Journal*, *5,* 251-279. https://doi.org/10.1007/s40797-019-00099-8.

Lee, B. S., Hong, S. H., & Wohar, M. E. (2017). City Size, Labor Productivity, and Wages In Korea. *The Singapore Economic Review*, *65*(4), 1073-1098. https://doi.org/10.1142/S0217590817500138.

Manning, C. (1980). Fringe Benefits in Manufacturing: Efficiency or Welfare. *Bulletin of Indonesian Economic Studies*, *16*(2), 54-82.

Matano, A., & Naticchioni, P. (2012). What Drives the Urban Wage Premium? Evidence. *Journal of Regional Science, 56*(2), 191-209. https://doi.org/10.1111/jors.12235.

Melo, P., Graham, D., & Noland, R. (2009). A Meta-analysis of Estimates of Urban Agglomeration Economies. *Regional Science and Urban Economics*, *39*(3), 332-342.

Nordman, C., Robilliard, A.-S., & Roubaud, F. (2011). Gender and Ethnic Earnings Gaps in Seven West African Cities. *Labour Economics*, *18*(S1), 132-145.

Nunn, N., & Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. *The Review of Economics and Statistics*, *94*(1), 20-36.

O'Clery, N., & Lora, E. (2016). City Size, Distance, and Formal Employmeent. *CID Research Fellow and Graduate Student Working Paper*, 1-17.

Peng, S. (2019). Urban Scale and Wage Premium: Evidence from China. *Journal of the Asia Pacific Economy*, *24*(3), 468-480.

Puga, D. (2010). The Magnitude and Cause of Agglomeration Economies. *Journal of Regional Science*, *50*(1), 203-219.

Roca, J. D., & Puga, D. (2017). Learning by Working in Big Cities. *Review of Economic Studies*, *84*(1), 106-142.

Rose, A., & Bright, E. (2014). The Landscan Global Population Distribution Project: Current State of the Art and Prospective Innovation. *Technical Report Oak Ridge National Laboratory (ORNL).*

Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and Sources of Agglomeration Economies. In J. V. Thisse (Ed). *Handbook of Regional and Urban Economics* (pp. 2119-2171). North Holland: Elsevier.

Sohn, K. (2015). Gender Discrimination in Earnings in Indonesia: A Fuller Picture. *Bulletin of Indonesian Economic Studies*, *51*(1), 95-121.

Tadjoeddin, M. Z. (2015). Spatial Dimensions of Key Employment Outcomes in Indonesia. *Journal of Comparative Asian Development*, *14*(3), 466-509.

Uchida, H., & Nelson, A. (2009). *Agglomeration Index: Towards a New Measure of Urban Concentration*. Washington D.C: World Bank.

Widya, A. H., Hartono, D., Indraswari, K. D., & Setyonugroho, L. D. (2019). Population Concentration and Productivity in the Metropolitan Area: Evidence from Indonesia. *International Journal of Economics and Management*, *13*(2), 453-466.

Yankow, J. J. (2006). Why Do Cities Pay More? An Empirical Examination of Some Competing Theories of the Urban Wage Premium. *Journal of Urban Economics*, *60*(2), 139-161. https://doi.org/10.1016/j.jue.2006.03.004.

Appendix 1. Variable Definition

No	Variable	Definition	Operational variable	Data category	Sources
(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables					
1	Wage per hour (wage_hr)	Wages per hour received by individuals		continuous data	Sakernas 2010 and 2015
2	Work	Individual during the last week working or not		1 work 0 tidak	Sakernas 2010 and 2015
3	Formal	Individuals work in the formal or informal sector	The formal informal sector is determined by status and type of work	1 formal 0 informal	Sakernas 2010 and 2015
4	city diversity (diver_city)	A measure of industrial diversity in a city		continuous data	Sakernas 2010 and 2015
Independent Variables					
5	Population density (density)	Total population per km2 of the district where the individual works		continuous data	SP 2010, Supas 2015, BPS
Individual Characteristics					
6	Age	Individual age at enumeration survey		continuous data	Sakernas 2010 and 2015
7	Squared age (age_sq)	The age of the individual at the time of enumeration is squared		continuous data	Sakernas 2010 and 2015
8	(experience)	The length of work on the job now		continuous data	Sakernas 2010 and 2015
9	Lama kerja kuadrat (exper_sq)	The length of work on the job now is squared		continuous data	Sakernas 2010 and 2015
10	Years of schooling (yos)		Years of schooling based on education completed Elementary school/ equivalent = 6 Middle school/ equivalent = 9 High School/ Equivalent = 12 Diploma I = 13 Diploma III = 15 Undergraduate = 16 postgraduate = 18	continuous data	Sakernas 2010 and 2015
11	Gender	Individual gender		1 male 0 female	Sakernas 2010 and 2015

http://journal.uinjkt.ac.id/index.php/signifikan
https://doi.org/10.15408/sjie.v12i1.27999

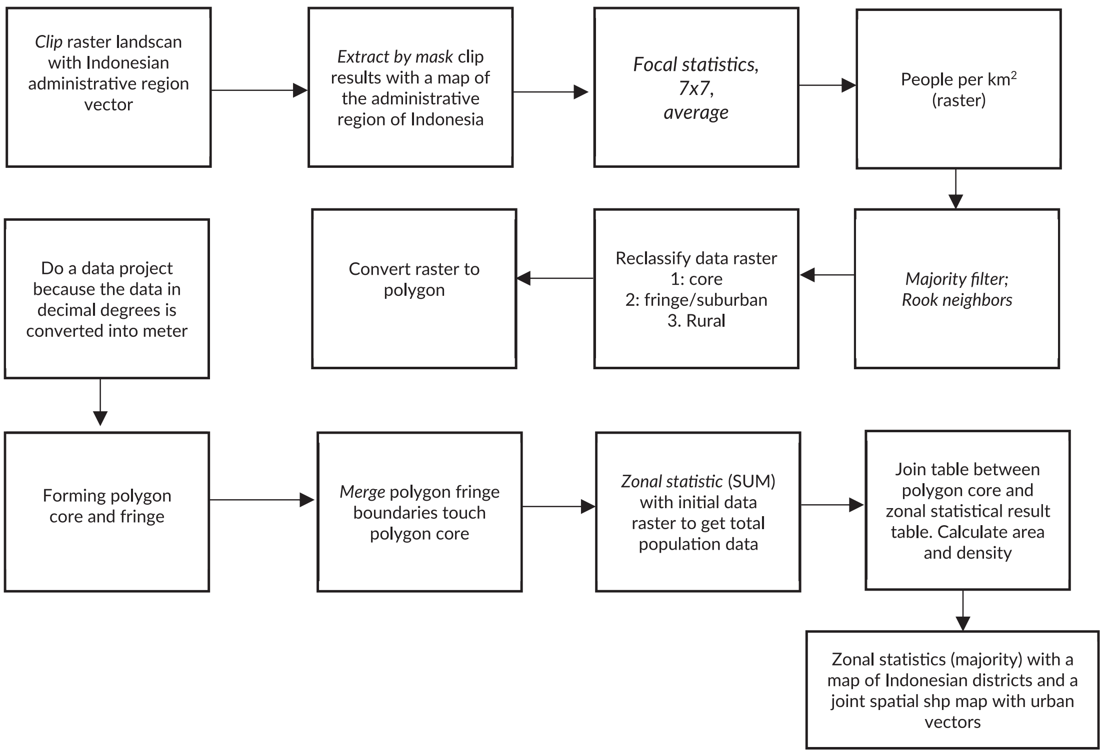
No	Variable	Definition	Operational variable	Data category	Sources
(1)	(2)	(3)	(4)	(5)	(6)
12	total working hours (workinghours)	total working hours for one week		continuous data	Sakernas 2010 and 2015
13	total working hours squared (workinghours_sq)	total working hours for one week is squared		continuous data	Sakernas 2010 and 2015
14	KBLI	individual business classification at current job (KBLI 2005)		Kode KBLI 2 digit	Sakernas 2010 and 2015
15	KJI	Type of job/ individual position at work (KJI 2002)		Based on Indonesian Job Classification (KJI)	Sakernas 2010 and 2015
Urban characteristics					
16	Human capital city (hc_city)	Share residents with high school education and above	Hc_city = the number of residents with high school education and above / the number of residents of the city/ district	continuous data	Sakernas 2010 and 2015
17	Island dummy (dpulau)	Individual work island code		1. Sumatera 2. Jawa 3. Bali 4. Nusa Tenggara 5. Kalimantan 6. Sulawesi 7. Maluku 8. Papua	Sakernas 2010 and 2015
Instrumental Variable					
18	Earthquake Disaster Risk Index (IRBI)	The level of disaster risk for each district/city in Indonesia is by the hazard it has		continuous data	Publication of Indonesia's Disaster Risk Index 2013
19	ruggedness	The average value of the terrain ruggedness at each center point in the Indonesian district. Terrain ruggedness is the absolute difference in the height of the center point with respect to 8 other adjacent points	ruggednes on the central cell grid i is the elevation on the grid of central cells i is the elevation on the grid of neighboring cells j Then a raster of index values is generated, which is then carried out by zonal statistics (means) with a map of Indonesian districts	continuous data	Geospatial Information Agency

Appendix 2. Constructing urban density variable

It was defining urban areas using ArcMap 10.5 software. The initial step taken is to determine the built cover threshold that matches the current commuting in Indonesia. After that, a smoothing algorithm is performed, the value of each reference cell is the average density of 7x7 cells around it, with the reference cell as its center. This step is done to overcome holes in urban areas, such as terrain, airports, or massive open public places. Population grids that meet the population density threshold per kilometer are grouped for adjacent grids. They are based on four rook neighbors (which do not include neighbors on the diagonal side), not based on eight queen neighbors (which have neighbors on the diagonal side).

Furthermore, it is classified into core, fringe, town, and rural. The threshold used in defining urban and suburban is: (a) core, with a population density of more than 4000 people per km^2 and a total population of more than 400,000 people, (b) fringe, with a population density of more than 2000 people per km^2, (c) suburban, a standing area alone, a population density of more than 2000 people per km^2 and a total population of more than 200,000 people. The area is said to be urban if it consists of cores and fringes.

The size of cities formed, urban and suburban is based on a population grid of 1 km^2. While other available data, such as wages and individual characteristics, are only available at the district/city level, there is no personal location point. Therefore, the size of the resulting city needs to be mapped to Indonesia's districts/cities. The way to do this is that a district belongs to an urban or suburban area, if at least 50 percent of the district's population is included in that urban area.



http://journal.uinjkt.ac.id/index.php/signifikan
https://doi.org/10.15408/sjie.v12i1.27999