Int. Journal of Economics and Management 13 (2): 453-466 (2019)

IJEM
International Journal of Economics and Management
Journal homepage: http://www.ijem.upm.edu.my

Population Concentration and Productivity in the Metropolitan Area: Evidence from Indonesia

ADISTA HANIF BASKARA WIDYA[a], DJONI HARTONO[a*] KENNY DEVITA INDRASWARI[a] AND LOURENTIUS DIMAS SETYONUGROHO[a]

[a]*Faculty of Economics and Business, Universitas Indonesia, Indonesia*

ABSTRACT

Economic activities are highly concentrated in a tiny geographical areas which are considered as areas providing increased returns. A vast amount of empirical evidence has shown that more significant population enhances urban agglomeration externalities, but small knowledge whether the urban population distribution influences that context. This study aims to examine how the urban population concentration in 10 Indonesian metropolitan areas affects productivity. Based on the estimation of a pooled cross-section time-series model from 2000 to 2014, this study revealed that the population distribution has a strong influence on the productivity of metropolitan areas. In terms of elasticity, an average increase of 1% in the degree of population concentration resulted in a rise in productivity per worker by 0.17%. However, this study also found that the combination of population concentration and population size might bring unintended effect within Indonesia's metropolitan areas setting. In summary, this study provided new evidence that explicitly expressed the fact that the degree of population concentration is a vital source of urban aggregate in agglomeration economies.

JEL Classification: O18, R11

Keywords: Agglomeration Economies; Metropolitan; Population Concentration; Productivity;

Article history:
Received: 17 February 2019
Accepted: 10 October 2019

[*] Corresponding author: Email: djoni.hartono@ui.ac.id

INTRODUCTION

The organization of economic activities in a dense area provides many benefits ranging from adequate amenities, easiness to find a match (employer, partner, supplier, and so forth), to ample opportunities of fruitful interaction with other economic agents. All these advantages, which would not exist without mass collection of economic activities in a geographical space, are externalities boosting productivity. The larger the agglomeration, the higher the productivity-enhancement effect. The fact that productivity is higher in larger cities was first noted by Smith (1776), Marshal (1890), and has been confirmed by modern empirical research. The vast empirical literature concluded that the higher the agglomeration intensity notated as city population size or average density, the higher the magnitude of the agglomeration externalities which is usually expressed as productivity per worker.

There is a need to surpass the standard approach and count for the spatial-structure dimension. It is then necessary to fully understand the nature of agglomeration externalities covering a metropolitan area. Despite the results in European and North American countries (Meijers and Burger 2010; Veneri and Burgalassi 2012), studies under this topic are still limited, particularly in Indonesia. Most empirical works on agglomeration economies do not go beyond the inclusion of the average density or city size as a spatially relevant factor. Consequently, the foremost objective of this study is to extend the understanding regarding how urban agglomeration economies can take place. Specifically, this study intends to empirically investigate the influence of the population distribution in the metropolitan areas on agglomeration externalities. Also, it aims to examine whether the impact of population distribution is augmented or not in increasing return from the population size.

The population distribution may affect the productivity of the metropolitan region in several ways. First, according to Krugman (1991), transportation cost is an important source of increasing returns. A high citizenry congregation can lead to have increasing returns on outputs in the production technology within a given geographical area. As transportation of products from one production stage to the next one necessitate costs, the distance is minimized by the proximity (Ciccone and Hall 1996). The second source is from lower commuting cost which is promoted by individual being closer to each other (Duranton and Puga 2004). It could enhance productivity since it allows the urban area to allocate resources to more productive activities other than commuting. As for industry, reduction of production costs and the improvement of innovation capacity and, further, to advancing total factor productivity (Han, Xie, and Fang 2018). Third, higher population concentration enables interaction between dwellers in metropolitan areas. From a microeconomic perspective, higher density level offers more significant opportunities for learning. Close physical proximity between urban dwellers helps the spread of knowledge by increasing the interaction and face-to-face contact (Abel, Dey, and Gabe 2011). One of the famous example is Silicon Valley where agglomeration advantages of manufacture and IT firms leading to an environment of mutual learning and assistance (Wetwitoo and Kato 2018; Börjesson et al. 2019). Those interlinkages have been shown to enhance productivity when information is not perfect with rapid alternation or difficulties in systemizing. These are the crucial attributes of the most inestimable economic activities presently (Storper and Venables 2004).

Although the idea of the concentrated urban region having advantages from the agglomeration was implicit in earlier pieces of literature, it seems that there was no early work in the Indonesian context. In the mentioned context, the population distribution was not an explicit element of the theory; there has not been empirical work based on the measurement of the population distribution. Therefore, this study aims to fill this gap. It presented a model of urban productivity which has been developed in several previous studies (Mills 1967; Moomaw 1983; Ciccone and Hall 1996; Abel, Dey, and Gabe 2011). This model explicitly incorporates population distribution and also the complementarity between population size and population distribution.

This paper is structured as follows — section 2 reviews some literature related to the scopes, sources, and measurement of agglomeration economies. Section 3 discusses the research methods and introduces the data used in the paper. Section 4 displays and analyzes the empirical findings. The last section outlines the conclusions and recommendations of the study.

LITERATURE REVIEW

At the broadest level, agglomeration economies are external economies emerging when individuals and firms are located near one another. It signifies that the productivity of workers or companies will increase with the overall amount of activity in other adjacent companies, or with the number of adjacent workers or consumers. The external economies subsist when the level of the suburban surroundings expands its productivity. There are no less than three aspects over which these externalities may extend.

Rosenthal and Strange (2004) referred to the industrial scope of the externality as the most familiar scope. The latter scope is referred to as the level to which agglomeration economies expand over industries, even beyond all sectors in a given city, rather than being enclosed within the diving lines of manufacturing.

The second scope deals with the geographical aspect. The aspect of significant geographic distance is the ease of agglomeration economies with separation: if agents are physically neighboring, then there will be a high potential for interchange. The third scope deals with the time-related aspect. Indeed, it is not impossible that an agent's conveyance with other agents at a given point long ago continues to influence current productivity. Thus, in addition to the rationally well familiar unchanged agglomeration economies, there may also be constantly changing agglomeration economies (Scott and Storper 2015).

Studies by Glaeser et al. (1992) and Henderson, Kuncoro, and Turner (1995) pointed out that the characteristics of a city can impact its growth throughout twenty years or more. However, it does not necessarily indicate that the economic environment, twenty years or so, pursues having a direct impact on growth. Alternatively, the impacts might be incidental through the accumulation of smaller straight effects over the period.

In terms of sources of urban agglomeration economies, Duranton and Puga (2004) distinguished three mechanisms through which urban increasing return can occur, namely sharing, matching, and learning. The mass sharing occurs when companies or employees profit by taking advantage of mutual resources while planning their activities. Another way in which mass economies may arise is through sharing which takes place when companies have the opportunity to share the gains from accessing a greater variety of intermediate inputs. Using the model of urban framework first developed by Abdel-Rahman and Fujita (1990), Duranton and Puga (2004) showed that the increase in the labor input of any particular sector in the city with perfectly competitive market structure, must be associated with more intermediate producers and also final producers that will become more productive when they have access to a broader range of varieties.

The next mechanism that can lead to agglomeration economies is matching. This notion views large cities as facilitators of the chances for matching. Employees will waste little time in seeking for jobs, leading to shorter unemployment. The reductions in these frictional costs can be used to feed lower labor costs or more output. (Overman and Puga 2009).

In addition, learning in a broad sense is a crucial parameter both in terms of resources devoted to it and its contribution to economic development. A fundamental feature of learning is that it is not a solitary activity taking place in a void (Duranton and Puga 2004) in most cases. Instead, it involves interactions with others, and many of these interactions have a 'face-to-face' nature, especially when it comes to tacit knowledge. Thus, even with modern communication technologies, distance still acts as a barrier to learning (Overman and Puga 2009). Towns may help education by gathering a tremendous number of people. The conveniences of those towns in education pay attention to cutting-edge technologies, acquisition of skills, daily incremental knowledge creation, and the diffusion and accumulation (knowing how, knowing who, etc.), as suggested by Lucas (1988), Rosenthal and Strange (2004), and Liang and Lu (2017).

To sum, there are lots of different ways in which physical proximity may aid sharing, matching or learning, and as a consequence, lead to mass economies. There was no difference between these mechanism based on the empirical analysis of the present study. The study instead focused on the overall impact on productivity.

Concerning measurement, the most common approach to observe the existence and magnitude of agglomeration economies is the production function aggregated to industrial levels within cities or metropolitan area levels as a whole (Meijers and Burger 2010; Sun,Wang, and Cai 2015). This approach was first developed by Mills (1967) as a general equilibrium model for an urban economy. The base form is as follow:

$$Q_i = g(A_i)f(X_i) \quad (1)$$

Where Qi represents the output of firm i or city i, Xi represents a vector of production inputs which usually consists of labor, capital stock, human capital, and other inputs. If the production function is assumed to exhibit constant returns to scale, the productivity-enhancement effect of agglomeration enters the production function through $g(Ai)$ that is known as a shift term. This approach is referred to as Hicks-neutral technical change production function. The second approach is to assume the production function to exhibit increasing returns to scale which originated from the agglomeration economies. If production input consists of labor and capital $Q_i = g(A_i)K_i^{\alpha}L_i^{\beta}$, the economies of agglomeration are measured by the summation of the coefficients on the inputs $\delta = \alpha + \beta$. This approach is known as a constant elasticity of substitution (CES) production function. The most popular approach that was also used in this study was the Hicks-neutral technical change production function. The reason was that the nature of the data used was good at controlling the return to the scale of the metropolitan area's production function. Also, Metropolitan Areas in Indonesia are characterized by a large difference in the composition of production input.

The shift term as a function of economies of agglomeration can be expressed to discuss how to capture the economies of agglomeration effect through shift term:

$$A_i = a + b\ AgglomerationScale_i + \varepsilon_i \quad (2)$$

The usual approach to a select variable as a surrogate of agglomeration economies is the shift term as a function of industry employment to capture localization economies, or urban employment/population size/ population density to capture urbanization economies. A given approach depends on the type of agglomeration scope being modeled, but for the urban area as a whole, there is no reason to focus on just one of the three approaches as they can all exist simultaneously (Eberts and Mcmillen 1999). However, there is also a possibility to employ a more general approach which incorporates many factors affecting productivity through shift term (Carlino 1992). The shift term can be expressed as an exponential function of the number of variables, as follows:

$$A_i = \exp[a_0 + \sum_{k=1}^{K} a_k\, x_{ki}] \quad (3)$$

Utilizing cross-section of the MSAs and a group of industries, Shefer (1973) concluded that doubling the size of the city would increase productivity by between 14-27%, whereas Sveikauskas (1975) revealed a productivity increment of only 6-7%. In addition, Segal (1976) demonstrated that productivity was roughly 8% higher in cities with populations of two million or more through revising the capital stock estimates of earlier research.

From a geographical scope front, Ciccone and Hall (1996), Harris and Ioannides (2000) tested the impact of labor density on productivity. Based on the US data at the state or metropolitan level, they found a positive effect of density on productivity with a doubling associated with roughly a 5% increase in productivity. From the European statistics, Ciccone (2002) revealed effects that were slightly inconsequential than those of the US. He further explained that the elasticity the workforce productivity regarding employment density was 4.5% in Europe, contrasted to 5% in the US. In the case of Japan, rents were utilized by Dekle and Eaton (1999) to consider mass economies using Japanese data at the prefecture level. Consequently, they found confirmations of mass economies in finance and manufacturing, although the magnitude was roughly one-quarter of the Ciccone and Hall (1996) estimates. Besides, it was also demonstrated that an increase in activity across Japan would raise the productivity of a given prefecture, suggesting a sizeable geographical scope.

Most empirical studies in recent years found similar results regarding the relationship between urban agglomeration and its economic benefits. Fujita and Thisse (2002) demonstrated that knowledge spillover from economic activity and population concentration was a strong force providing a trickle-down effect in the urban community. Similarly, Martin and Ottaviano (2001) and Li and Liu (2018) supported the result that economic productivity was significantly associated with the spatial structure, where a higher degree of agglomeration provided high economic productivity. However, some empirical analysis disproved that notion. In an investigation of urban concentration in 68 countries from 1985-2010, Frick and Rodriguez-Pose (2017) suggested that there was no consistent relationship between urban density and economic growth as it varied between developed and developing countries. Other studies, such as Meijers and Burgers (2010) in the US and Veneri and Burgalassi (2012) in Italy, found a positive relationship between polycentric and productivity levels. Thus, even though most studies suggested a positive relationship, the relationship between urban density and productivity was debatable depending on the situation of the country.

Finally, there are various methods used to measure population distribution, namely Hoover Index, entropy index, Gini index. This study will use the Hoover Index (Hoover 1936, 1941) as it is the most widely used measure for assessing the concentration or de-concentration tendencies of the evolving population distribution of a region[1]. This measure can be calculated through the following formula:

$$H_{m,t} = 0.5 \sum_{i=0}^{n} \left| \frac{u_{i,t}}{U} - \frac{a_{i,t}}{A} \right| 100 \tag{4}$$

Where, H is the metropolitan Hoover Index, m indicates metropolitan area, t represents time, n indicates the total number of city comprising a metropolitan area, u_i is the metropolitan area's population residing in city i, U is the total metropolitan area's population, a_i is the city i total area (km^2), and A is the total metropolitan region's area (km^2). The index's value ranged from 0 to 100. If the population was equally distributed within the metropolitan region, then the index would be approaching zero, indicating high levels of population dispersal. The index value would be close to 100 if the population in the metropolitan region were highly concentrated in a single city. Thus, the high value of the Hoover Index might be associated with high agglomeration externalities.

METHODOLOGY

This study presented a general model of urban production function built on previous work to construct a conceptual framework for the empirical analysis (Mills 1967; Moomaw 1983; Ciccone & Hall 1996; Abel, Dey, and Gabe 2011). The production function was aggregated to the metropolitan area level:

$$\mathrm{Q}_{it} = A(.)_{it} {K_{it}}^{\alpha} {H^{\beta}}_{it} {G_{it}}^{\theta} {N_{it}}^{1-\alpha-\beta-\theta} \tag{5}$$

Q represents output, K denotes physical capital, H is human capital, N is the number of workers employed, and A (.) represents Hicks-neutral technology parameter acting as the shift term of the production function. Subscript i and t represent the metropolitan area and time, respectively. The measure of infrastructure capital stock (G) was also included as a production input. The reason was that urban public infrastructure might directly affect the operation efficiency of the cities, particularly large cities. As a result, it promotes the realization of agglomeration economies (Eberts and Mcmillen 1999).

The parameter α,β,θ, and 1-α-β-θ represent the elasticity of output concerning physical capital, human capital, infrastructure capital, and labor. The assumption that the production function exhibits a constant return to scale was invoked, i.e., α+β+θ=1. Thus, the productivity-enhancement effect of agglomeration entered the production function through the Hicks-neutral technology parameter A(.). Furthermore, the metropolitan area's productivity per-worker can be expressed as in a linear stochastic form as:

[1]. The index is considered to be a standard to evaluate the fairness of dispersal of the population within a given geographical location. It is built on an index and is meant to represent the Gini version of diversity index (index dissimilarities). It compares the percentage of the population of each city with a proportional share of the total population of metropolitan areas. It also measures the correspondence degree between the population and territory (Santic, 2011).

$$\ln\left(\frac{Q}{N}\right)_{it} = lnA(.)_{it} + \alpha\, ln\left(\frac{K}{N}\right)_{it} + \beta\, ln\left(\frac{H}{N}\right)_{it} + \theta\, ln\left(\frac{G}{N}\right)_{it} + v_{it} \tag{6}$$

According to Carlino and Voith (1992), to examine the production function's shift term, the shift term can be expressed as a function of the number of variables affecting productivity through it:

$$A_{it} = exp[\gamma_0 + \textstyle\sum_{k=1}^{K} \gamma_k x_{kit}] \tag{7}$$

As it was hypothesized that agglomeration economies were jointly determined by the metropolitan area's population size and its distribution, the expression of the shift term in multiplicative form was as follows:

$$\mathrm{A}_{it} = \gamma_0\, P_{it}^{\gamma_1}\, H_{it}^{\gamma_2} \tag{8}$$

P is the population size, and H is the population distribution as measured by the Hoover Index. High Hoover Index value indicates a more concentrated urban area. γ_1 and γ_2 represent the elasticity of output concerning population size and population distribution. γ_0 denotes other factors of the technology parameter that is independent of population size and its distribution. Importantly, the parameter γ_1 and γ_2 measure the net agglomeration effect of population size and population distribution, incorporating both the positive and negative (congestion) effects arising from the measures. Thus, this parameter will depend on the relative strength of each opposing force. So, the technology parameter can be expressed in linear stochastic form as:

$$ln\,(\mathrm{A})_{it} = \gamma_0 + \gamma_1\, lnP_{it} + \gamma_2\, ln\, H_{it} + \gamma_3 INDSTRC_{it} + v_{it} \tag{9}$$

INDSTRC indicates the metropolitan area's industrial share as a control variable accounting for the level of the metropolitan area's industrialization (Carlino & Voith 1992). The variable was the share of manufacturing activity over the GDP. Moreover, by substituting equation (9) to equation (6):

$$ln\left(\frac{Q}{N}\right)_{it} = \gamma_0 + \gamma_1\, lnP_{it} + \gamma_2\, ln\, H_{it} + \alpha\, ln\left(\frac{K}{N}\right)_{it} + \beta\, ln\left(\frac{H}{N}\right)_{it} + \theta\, ln\left(\frac{G}{N}\right)_{it} + \gamma_3 INDSTRC_{it} + w_{it} \tag{10}$$

Equation (10) serves as a baseline model for the empirical analysis. Even equation (10) can validate whether population distribution has any impact on the metropolitan area's productivity or not; the specification does not allow for the interaction of population size and population distribution. Meanwhile, a high concentration of population can enhance the increasing return from population size by increasing the ability of the metropolitan area's dwellers to share gains from indivisible facilities and higher chances of matching worker and employer. The research was also interested in accounting for the possibility of the empirical estimation model. Following the approach that Abel, Dey, and Gabe (2011) employed in examining the complementarity between population density and human capital, the complementarity between population size and population distribution was formally expressed by assuming the elasticity of output concerning population size variations with the degree of population concentration as follows:

$$\gamma_1 = \delta_0 + \delta_1 \ln H_{ij} \qquad\qquad \delta_1 > 0 \tag{11}$$

δ_1 represents the contribution of population concentration to the net effect of population size while δ_0 denotes other factors of this parameter that are independent of population concentration. It was assumed that $\delta_1 > 0$ implies that the population distribution and population size were complementary in production. Substituting equation (11) to equation (10) gave the second estimating model as follows:

$$ln\left(\frac{Q}{N}\right)_{it} = \gamma_0 + \delta_0\, lnP_{it} + \delta_1 (lnP_{it})(lnH_{it}) + \gamma_2\, ln\, H_{it} + \alpha\, ln\left(\frac{K}{N}\right)_{it} + \beta\, ln\left(\frac{H}{N}\right)_{it} + \theta\, ln\left(\frac{G}{N}\right)_{it} + \gamma_3 INDSTRC_{it} + e_{it} \tag{12}$$

Equation (10) provides the tool to examine whether or not the population distribution is influencing in determining the magnitude of the metropolitan area's agglomeration economies, while equation (12) provides the foundation to gain more insight on how exactly population distribution influences productivity. Specifically, equation (12) might explain if a higher concentration of population enhances the productivity-enhancement effect from population size.

Before turning into the evidence of this study, it is important to note that by adding interaction variable into equation (10), the study was only interested in the significance of the interaction term rather than interpreting the value and significance of the term used to compute them (Williams 2015). The technique employed was similar to how Abel, Dey, and Gabe (2011) analyzed the interaction between urban density and human capital in influencing productivity. They ignored the results of their main variables that are density and human capital, which yield negative value when the interaction term was entered and only interpreting the significance of their interaction variable.

Moreover, if the first derivative of the equation is taken (12) concerning population size, it might lead to the impact of population size to productivity per worker which increases with the degree of metropolitan area's population concentration. The equation is as follows:

$$\frac{\delta[ln(Q_{it}/N_{it})]}{\delta[\ln(P_{it})]} = \delta_0 + \delta_1(lnHI_{it}) \quad (13)$$

If $\delta_1$is positive and significant, means that the net impact of population size to metropolitan area's productivity per worker is a function of a constant (δ_0) and a coefficient (δ_1), the value of which increases with the increase in the Hoover Index ($lnHI_{it}$).

Before performing the econometric estimation, some estimation issues need to be addressed. First, the metropolitan area is defined as composed by a set of core cities and several satellite cities and regencies surrounding the core area. However, since the data on all the variables are not available at the metropolitan area levels, the aggregation was carried out manually. Second, one of the most challenging aspects in estimating agglomeration economies dealt with endogeneity problem which led to overestimating the relationship between agglomeration and output (Rosenthal & Strange 2003). The 10-year lag was considered as an appropriate measure to obtain robust estimation. In other words, in the estimation model, the metropolitan area's productivity in the year 2000 was influenced by its population size and population distribution in the year 1990. This approach could be seen as an instrumental variable approach being used by Glaeser et al. (1992), Henderson, Kuncoro, and Turner (1995), and Combes et al. (2010). They argued that this approach was well-grounded as agglomeration economies are dynamic forces in the sense that the characters of a city can impact its growth throughout twenty years or more. Thus, the baseline empirical estimation model became:

$$n\,(\frac{\mathrm{Q}}{\mathrm{N}})_{it} = \gamma_0 + \gamma_1\, lnP_{it-10} + \gamma_2\, ln\, H_{it-10} + \alpha\, ln(\frac{K}{N})_{it} + \beta\, ln(\frac{H}{N})_{it} + \theta\, ln(\frac{G}{N})_{it} + \gamma_3 INDSTRC_{it} + \varepsilon_{it} \quad (14)$$

Third, since the data on capital stock and infrastructure capital are practically not available at all levels, the data were estimated by using the perpetual inventory method[2]. The basic idea of the perpetual inventory method (PIM) is that the net capital stock at the beginning of the period *t*, (K_t), is a function of the net capital stock at the beginning of the previous period *t-1*, (K_{t-1}), gross investment in the previous period, (I_{t-1}), and depreciation rate (δ):

$$K_t = (1-\delta)K_{t-1} + I_t \quad (15)$$

The previous period's capital stock was estimated (K_{t-1}) from the initial investment (I_t), the long-term growth rate of investment, (g_i), and the use of 10% rate of capital depreciation (δ), which was in line with Schundeln's (2007) research that found that the depreciation rates of physical capital invested in manufacturing enterprises in Indonesia. This estimation was built based on the establishment-level survey data to be between 8% and 12%. The formula to obtain K_{t-1}, was as follows:

$$K_{t-1} = \frac{I_t}{g_I + \delta} \quad (16)$$

[2]. The use of this method in the area of the urban economy can be found in (Wu, 1999) and (Xu, 2009).

The long-term growth rate (g_i) and the initial investment (I_t) value were calculated from a regression approach. This study used the time series data set of gross fixed capital formation data available and regressed the time series of log investment for province *i* on time *t*. Thus, the equation was estimated as follows:

$$\ln I_{i,t} = a_i + b_i t + \varepsilon_{i,t} \tag{17}$$

Where the transformed value of a_i acts as the initial investment and b_i as the long-term growth rate.

The samples of this study were ten metropolitan areas in Indonesia. The observation periods were the productivity from the years 2000 to 2014, where productivities were influenced by population size and population distribution from the years 1990 to2004. The data of the years 1991 to 1994 were absent, providing 110 observations. The observed periods were chosen because they were the transition periods for Indonesia departing from the pre-reformation to post-reformation era. Thus, it is a suitable period to estimate the dynamic shift that happens in urban areas in Indonesia. The dependent variable of the estimation model, labor productivity, was calculated by dividing the real metropolitan area's GDP by its total employment. Human capital was proxied by the number of workers in the metropolitan area having at least a diploma degree. Data on capital stock and infrastructure capital were obtained by the perpetual inventory method. The industrial structure was obtained by dividing the metropolitan area's GDP of the manufacturing sector by its total GDP. Hoover Index and Population Size were based on Census data in the years 1990 to 2004 aggregated according to each metropolitan area definition. Capital per worker was calculated from gross fixed capital formation data from the World Bank. Data of total employment and human capital were obtained from labor force survey. Moreover, data from metropolitan regions' GDP were obtained from the World Bank.

RESULTS

Data Description

In a correlation test, Figure 1 indicated that population size was positively correlated with the metropolitan area's productivity. Metropolitan area with the high population was associated (not necessarily causal relationship) with a high level of productivity per worker. Preliminary observation also suggested that metropolitan areas with a high concentration of population were associated with high levels of productivity as depicted in Figure 2. A more interesting finding was that the slope of the regression line in Figure 2 was sharper than the slope of the regression line in Figure 1. If the causal correlation depicted in Figure 1 and Figure 2 was statistically correct, the increase in the population concentration of a metropolitan area by 1% was more beneficial for productivity per worker than a 1% increase in population size. However, regression analysis should be conducted to determine further whether the samples in the metropolitan area’s productivity necessarily influence their population size and degree of concentration or not.

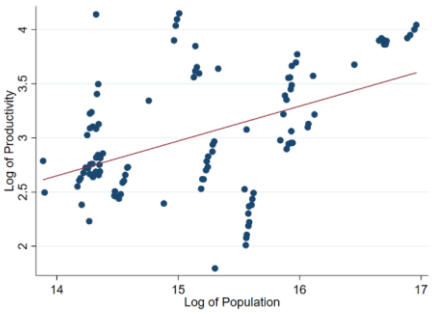

Figure 1 Productivity (2000 – 2014) and City Population Size (1990 – 2004) in 10 Metropolitan Areas

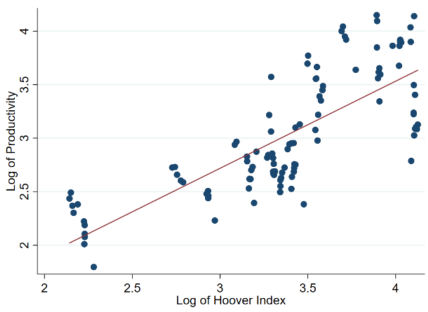

Figure 2 Productivity (2000 – 2014) and Population Distribution (1990 – 2004) in 10 Metropolitan Areas

Based on Appendix 2, the most concentrated metropolitan areas in the study sample was Makassar, whereas Surakarta was the least concentrated[3]. The fact that Makassar was the most concentrated metropolitan area was quite interesting, given that it was the least populated area. Jakarta with the highest number of dwellers in the samples was just below Makassar with a Hoover Index value of 55.65 in 1990.

In terms of productivity differences across Indonesia Metropolitan Areas, Table 1 presented a ranking of 10 Indonesian metropolitan areas based on average output per worker, population size, and degree of population concentration. On average, Jakarta having the highest number of the population turned out to be the most productive metropolitan area in the samples. The presence of Surabaya among the top three metropolitan areas with the highest productivity per worker further supported that argument as it ranked in the third place as the most crowded urban area.

Table 1 Average Output per Worker, Population Size, and Population Concentration for 10 Indonesia Metropolitan Areas, 1990 - 2014

Metropolitan Area (MA)	Output Per Worker (Million IDR)		Population Size		Hoover Index	
	Average	Rank	Average	Rank	Average	Rank
Jakarta MA	47.95	1	19,090,884	1	50.21	3
Medan MA	40.48	2	3,537,394	6	50.75	2
Surabaya MA	30.52	3	7,987,659	3	34.85	4
Makassar MA	24.99	4	1,578,971	9	60.85	1
Bandung MA	19.91	5	8,474,541	2	29.82	5
Semarang MA	14.58	6	4,046,338	5	23.57	8
Banjarmasin MA	14.53	7	1,589,430	8	27.82	7
Denpasar MA	13.95	8	1,571,270	10	29.18	6
Yogyakarta MA	12.21	9	1,999,578	7	17.46	9
Surakarta MA	8.76	10	5,729,837	4	9.05	10

Interestingly, Medan and Makassar ranked among the lowest areas in terms of population size and occupied the second and fourth places in terms of productivity per worker, respectively. This outcome might be explained by the degrees of their population concentrations. Based on Table 1, Makassar and Medan, even if ranked among the lowest in terms of population size, performed very well in the domain of population concentration. This finding further strengthened the validity of the hypothesis of the study as their populations were relatively concentrated compared to those of other metropolitan areas. They were able to realize high productivity gain from a small population size.

Estimation Results

Table 2 reports the estimation results for seven versions of equation. The results were based on the estimations of a pooled cross-section time-series model for 10 metropolitan areas in Indonesia for the period 2000-2014 (providing 110 observations). The results of the Hausman test for estimation results (6) acting as the baseline model was 0.4226 (presented at the bottom of Table 2), leading to the rejection of the hypothesis of non-systematic difference between fixed effect and random effect estimates. Thus, all the results presented in Table 2 were estimated using random effect estimation. The result of a global test of all the estimation models were below 5%, indicating that all the coefficients in each model were different to zero. The value of the R-Squared reported in the table for the baseline model also implicated that the baseline model was quite robust, explaining nearly 92 percent of the variation of output per worker across Indonesian Metropolitan Areas.

[3]. The Hoover Index value for Makassar was 59.65 in the year 1990. Makassar had 68% of its population living in the core city while the remaining 32% lived in the rural part whereas the Hoover Index value for Surakarta was 9.78 in the same year. Surakarta had only 9 percent of its dwellers residing in the core city.

Table 2 Regression Results

Variables	Ln(GDP/N), 2000 – 2014 (1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln(Hoover Index), 1990 - 2004	0.399*** (0.142)		0.661*** (0.0931)	0.177*** (0.0553)		0.169*** (0.0580)	2.929*** (1.017)
Ln(Population), 1990 - 2004		0.596*** (0.120)	0.372*** (0.0615)		-0.0609 (0.0562)	-0.0200 (0.0504)	0.618*** (0.240)
Interaction Term							-0.180*** (0.0661)
Ln(K/N), 2000 - 2014				0.689*** (0.0390)	0.752*** (0.0467)	0.700*** (0.0468)	0.705*** (0.0455)
Ln(H/N), 2000 - 2014				-0.161*** (0.0231)	-0.151*** (0.0226)	-0.159*** (0.0234)	-0.164*** (0.0227)
Ln(G/N), 2000 – 2014				0.0152 (0.0355)	-0.0356 (0.0355)	0.0103 (0.0373)	0.00483 (0.0363)
Industry Share, 2000 – 2014				-0.172 (0.328)	-0.0606 (0.364)	-0.140 (0.338)	-0.0901 (0.330)
Constant	1.616*** (0.490)	-6.043*** (1.824)	1.616*** (0.490)	-0.172 (0.243)	1.120 (0.823)	0.113 (0.761)	-9.702*** (3.688)
Observations	110	110	110	110	110	110	110
Number of region1	10	10	10	10	10	10	10
Prob > Chi2	0.0578	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R2	0.605	0.221	0.605	0.919	0.849	0.916	0.927

Note: a. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. b. Chi-square value for the Hausman test of the model (7) was 7.06; failing to reject the hypothesis of non-systematic difference between fixed effects and random effects estimates. c. R2 values reported for these models were the overall R2 values. d. Using Wooldridge test for autocorrelation in panel data, we found panel autocorrelation in equation (6) (Prob>F = 0.0125). Hence, we apply RE GLS regression with AR(1) disturbances for all equation above.

This study relied on an instrumental variable approach in which the agglomeration variables were proxied by 10-year lags value to confirm the robustness of the results (see Glaeser et al. 1992; Henderson, Kuncoro, and Turner 1995; Combes et al. 2010). Based on the estimation results (1), the impact of population distribution on productivity was indeed statistically significant. Moreover, the estimation result (2) also confirmed that the correlation between population size and productivity in the sample was not a mere spurious relation. Significant effect between population distribution and productivity is also found in Li and Liu (2018) and Lee and Gordon (2007) study. Decreasing production costs and the improvement of innovation will advance total factor productivity, even in one empirical study, encourage energy efficiency (Han et al., 2018). The elasticity of the productivity regarding the population distribution was higher than that of productivity regarding the population size as shown in estimation results (3). However, the results presented by the estimation results (1) to (3) might be subject to severe biased due to the omitted variable issue, especially data on capital stock (Rosenthal and Strange, 2003). Thus, it shall be turned into examining the estimation results (4), (5), and (6), including metropolitan area's production inputs and industrial structure as control variables in the model.

Based on the estimation results (4) even without including population size into the model, the population distribution alone was still having a positive and significant impact on productivity. Interestingly, the estimation results (5) showed that by excluding population distribution from the model, the impact of population size towards productivity became unclear. This result was similar to that of Carlino (1978) and Duranton & Turner (2015), concluding that the population size was not an appropriate measure for agglomeration economies. That is because a higher number of population does not mean an efficient communication and production occure in a one place. Thus, some studies consider more on population density or efficient access compared to population size (Gerritse and Arribas-Bel 2018).

Model (6) examined jointly population size and population distribution impacts on productivity with a model that included the metropolitan area's production input and industrial structure as control variables. Similar with the finding in equation (4) and (5), equation (6) showed that population distribution had a positive and significant impact on increasing the metropolitan area's productivity, meanwhile the association between population size and productivity was still unclear. Table 2 shows that a 1% increase on Hoover Index was associated with 0.17% increase on productivity. This result fosters conclusion from previous equations that the elasticity of productivity was more positively connected with the dynamics on population distribution rather than population size.

To test the second hypothesis stating that an increase in return from the population was high when the metropolitan area was concentrated, a regression was run for equation (6) and included interaction term between population size and population distribution. It allowed the elasticity of productivity with respect to the population size to vary with the degree of metropolitan area's population concentration[4]. Estimation (7) showed that an interaction between population distribution and number of population had a significant and negative impact on productivity. It implied that the combination of highly concentrated population and large population size had hampered the realization of potential productivity gain from agglomeration. Severe traffic congestion is one of the example on how the combination between highly concentrated population and large population size brought a negative effect on productivity. This result is also in line with Melo et al. (2017) study in US metropolitan cities which even conclude that there is no economic argument for allocating more public money to improve infrastructure in metropolitan areas with 1 million people or more. Population agglomeration was shown to be an important factor for increasing total productivity in a regional economy. However, productive efficiency and immediate access to jobs is also important to promote regional productivity (Melo et al. 2017; Otsuka 2017). Thus, some studies highlights the importance of investing in efficient transport networks (Melo et al. 2017).

In summary, the overall findings exhibited that population concentration was having both direct influence on metropolitan area's productivity. The direct influence might emerge from the increasing returns from the technology of production which became apparent due to high concentration promoting minimization of transportation cost in production (Ciccone and Hall 1996). Second, high population concentration might allow the metropolitan area to perform better in productivity per worker since it also minimized commuting cost, allowing a metropolitan area to allocate more resources to productive economic activities (Duranton and Puga 2004). These direct benefits could be seen as increasing returns from metropolitan area's spatial structure.

CONCLUSION

There has been a resurgence in academic works providing evidence on how population size creates agglomeration externalities which in turn translated into productivity. However, little is known about how population de-concentration may affect metropolitan area agglomeration externalities. Motivated by this setting, using a model of urban productivity incorporating population distribution dimension and the complementarity between population size and the degree of population concentration, this study aimed to examine how the spatial structures of 10 Indonesian metropolitan areas, as indicated by the 'centralization-dispersion' dimension, might affect their productivity.

Based on the estimation results, the present study indicated that population distribution was indeed having a strong influence on the metropolitan area's productivity. On average, we found that the effect from population concentration, as indicated by Hoover Index, significantly increases the productivity of metropolitan area. However, we also found that the effect of population size on metropolitan area's productivity is unclear. Further, we also found that the combination of highly concentrated population and large population size brought a negative impact on productivity. To conclude, this research provided new evidence that explicitly expressed that the degree of population concentration was an essential determinant of the urban aggregate of the agglomeration economies, even though its combination with population size might bring unintended effect within Indonesia's metropolitan areas setting.

There were several limitations to this study. First, the 'centralization-dispersion' dimension of the metropolitan area might not perfectly be captured by the Hoover Index, since it was computed using total area instead of total built-up area. Second, the study explicitly examined the impact of population distribution on productivity. Nevertheless, another characteristic of the metropolitan area's spatial structure might also

[4] By adding interaction variable into equation (6) this study was only interested in the significance of the interaction term, rather than interpreting the value and significance of the term used to compute (Williams, 2015).

contribute to productivity such as the 'monocentric-polycentric.' Finally, the limitations of this study might provide further directions for future studies under the same topics.

REFERENCES

Abdel-Rahman, H.M. & Fujita, M. 1990, 'Product Variety, Marshallian Externalities and City Size', *Journal of Regional Science*, vol. 30, pp. 165-183.

Abel, J.R., Dey, I. & Gabe, T. M. 2011, 'Productivity and the Density of Human Capital', *Federal Reserve Bank of New York Staff Reports*, March 2010, Revised: September 2011 vol. 440, viewed 12 April.

Börjesson, M., Isacsson, G., Andersson, M., & Anderstig, C. 2019, 'Agglomeration, productivity and the role of transport system improvements', *Economics of Transportation*, vol. 18(June 2018), pp. 27–39. https://doi.org/10.1016/j.ecotra.2018.12.002

Carlino, G.A. 1978, *Economies of Scale in Manufacturing Location: Theory and Measurement*, Springer US, Boston.

Carlino, G. & Voith, R. 1992, 'Accounting for differences in aggregate state productivity', *Regional Science and Urban Economics*, vol. 22, pp. 597–617.

Ciccone, A. & Hall, R. 1996, 'Productivity and the Density of Economic Activity', *The American Economic Review*, vol. 86, pp. 55-70.

Ciccone, A. 2002, 'Agglomeration effects in Europe', *European Economic Review*, vol. 46, no. 2, pp. 213-227.

Dekle, R. & Eaton, J. 1999, 'Agglomeration and Land Rents: Evidence from the Prefectures', *Journal of Urban Economics*, vol. 46, pp. 200-214.

Duranton, G. & Puga, D. 2004, *Micro-Foundations of Urban Agglomeration Economies*', Handbook of Urban and Regional Economics, North Holland, Elsevier, New York.

Duranton, G. & Turner, M. A. 2015, *Urban form and driving: Evidence from US cities (Mimeo)*. Providence: Brown University, United States

Eberts, R. W. & Mcmillen, D. P. 1999, *Agglomeration economies and urban public infrastructure*, Handbook of Regional and Urban Economics, North Holland, Elsevier, New York.

Frick, S. A. & Rodríguez-Pose, A. 2017, 'Big or Small Cities? City Size and Economic Growth', *Growth and Change,* vol. 49, no. 1, pp. 4–32. https://doi.org/10.1111/grow.12232

Fujita, M. & Thisse, J. F. 2002, *Economics of Agglomeration: Cities, industrial location, and globalization*. Cambridge University Press, Cambridge.

Gerritse, M. & Arribas-Bel, D. 2018, 'Concrete agglomeration benefits: do roads improve urban connections or just attract more people?', *Regional Studies*, vol. 52, no. 8, pp. 1134–1149.

Glaeser, E., Kallal, H., Scheinkman, J. & Shleifer, A. 1992, 'Growth in Cities', *Journal of Political Economy*, vol. 100, pp. 1126-1152.

Han, F., Xie, R. & Fang, J. 2018, 'Urban agglomeration economies and industrial energy efficiency', *Energy*, vol. 162, pp. 45–59. https://doi.org/10.1016/j.energy.2018.07.163

Henderson, J. V., Kuncoro, A. & Turner, M. 1995, 'Industrial Development in Cities', *Journal of Political Economy*, vol. 103, pp. 1067-1085.

Hoover, E. M. Jr. 1936', 'The Measurement of Industrial Localization', *Review of Economics and Statistics*, vol. 18, no. 4, pp. 162–171.

Hoover, E. 1941, 'Interstate Redistribution of Population 1850–1940', *Journal of Economic History*, vol. 1, pp. 199–205.

Krugman, P. 1991, 'Increasing Returns and Economic Geography', *Journal of Political Economy*, vol. 99, no. 3, pp. 483-499.

Li, Y. & Liu, X. 2018, 'How Did Urban Polycentricity and Dispersion Affect Economic Productivity? A Case Study of 306 Chinese Cities', *Landscape and Urban Planning*, vol. 173, pp. 51-59.

Lee, B. and Gordon, P., 2007, February, Urban spatial structure and economic growth in US metropolitan areas, In 46th annual meetings of the western regional science association, at Newport Beach, CA.

Liang, W. & Lu, M. 2017, 'Growth led by Human Capital in Big Cities: Exploring complementarities and spatial agglomeration of the workforce with various skills', *China Economic Review,* vol. 57, pp. 101-113 http://dx.doi.org/10.1016/j.chieco.2017.09.012.

Lucas, Robert E., Jr. 1988, 'On the Mechanics of Economic Development', *Journal of Monetary Economics*, vol. 22, pp. 3-42.

Martin, P. & Ottaviano, G. I. P. 2001, 'Growth and agglomeration. International Economic Review', vol. 42, no. 4, pp. 947–968. https://doi.org/10.1111/1468-2354.00141

Meijers, E. J. & Burger, M. J. 2010, 'Spatial structure and productivity in US metropolitan areas', *Environment and Planning A: Economy and Space*, vol. 42, no. 6, pp. 1383–1402

Melo, P. C., Graham, D. J., Levinson, D. & Aarabi, S. 2017, 'Agglomeration, accessibility and productivity: Evidence for large metropolitan areas in the US', *Urban Studies*, vol. 54, no. 1, pp. 179–195. https://doi.org/10.1177/0042098015624850

Mills, E.S. 1967, 'An Aggregative Model of Resource Allocation in a Metropolitan Area', *American Economic Review,* vol. 57, pp. 197-210.

Moomaw, R.L. 1983, 'Is Population Scale a Worthless Surrogate for Business Agglomeration Economies?', *Regional Science and Urban Economics*, vol. 13, pp. 525-545.

Otsuka, A. 2017, 'Regional determinants of total factor productivity in Japan: stochastic frontier analysis', *The Annals of Regional Science*, vol. 58, 3, pp. 579–596. https://doi.org/10.1007/s00168-017-0808-7

Overman, H. G. & Puga, D. 2009, 'Labour Pooling as a Source of Agglomeration', *Agglomeration Economics*, pp. 133-150.

Rosenthal, S. S. & Strange, W. C. 2004, *Evidence on the nature and sources of agglomeration economies*, Handbook of Regional and Urban Economics, North Holland, Elsevier, New York.

Santic, D. 2014, 'The spatial concentration of population in Serbia 1981-2011 measured with the Hoover Index', *Proceedings of the Matica Srpska for Social Sciences*, vol. 148, pp. 461-470.

Scott, A. J. & Storper, M. 2015, 'The nature of cities: The scope and limits of urban theory', *International Journal of Urban and Regional Research*, vol. 39, no. 1, pp. 1–15.

Segal, D. 1976, 'Are There Returns to Scale in City Size?', *Review of Economics Statistics*, vol. 58, pp. 339-450.

Shefer, D. 1973, 'Localization Economies in SMSAs: A Production Function Analysis', *Journal of Regional Science*, vol. 13, pp. 55-64.

Smith, A., 1776. An inquiry into the nature and causes of the wealth of nations: Volume One. London: printed for W. Strahan; and T. Cadell, 1776.

Storper, Michael & Anthony J. Venables. 2004, 'Buzz: Face-to-Face Contact and the Urban Economy', *Journal of Economic Geography*, vol. 4, no. 4, pp. 351-370.

Sun, B., Wang, X. & Cai, Y. 2015, 'An empirical study of the economic performance of polycentric spatial structure of mega-cities in China', *City Planning Review*, vol. 39, no. 8, pp. 39–45.

Sveikauskas, L. 1975, 'The Productivity of Cities', *Quarterly Journal of Economics*, vol. 89, pp. 393-414.

Veneri, P. & Burgalassi, D. 2012, 'Questioning polycentric development and its effects. Issues of definition and measurement for the Italian NUTS-2 regions', *European Planning Studies*, vol. 20, no. 6, pp. 1017–1037.

Wetwitoo, J. & Kato, H. 2017, 'Inter-regional transportation and economic productivity: a case study of regional agglomeration economies in Japan', *The Annals of Regional Science*, vol. 59, no. 2, pp. 321–344. https://doi.org/10.1007/s00168-017-0833-6

Williams, Richard. 2015, 'Interaction Effects between Continuous Variables', *The University of Notre Dame*, February 20, 2015, viewed July 2018 < http://www3.nd.edu/~rwilliam/>.

APPENDIX

Appendix 1 Summary Statistics

Variable	Number of Obs.	Mean	Standard Deviation	Min.	Max.
Main Variables					
Output per Worker	150	22.79	13.87	6.02	63.38
Hoover Index	110	33.35	15.66	8.52	61.94
Population Size	110	5,560,590	5,263,366	1,073,616	23,200,000
Production Input					
Capital per Worker	150	31.97	28.26	6.00	142.44
Human Capital per Worker	150	0.09	0.04	0.03	0.25
Infrastructure Capital per Worker	150	0.80	0.56	0.22	3.36
Industry Share					
Manufacturing GDP/Total GDP	150	0.27	0.09	0.09	0.45

Notes: Output per worker is GDP/Total Employment from 2000 to 2014. Hoover Index and Population Size are based on Census data in the year of 1990 to 2004 aggregated according to each metropolitan area definition. Capital per worker is calculated from gross fixed capital formation data from INDODAPOER (World Bank) using the perpetual inventory method. Infrastructure capital is calculated from infrastructure expenditure function data from INDODAPOER (World Bank) using the perpetual inventory method. Human capital is proxied by the number of people (+15) with a four-year college degree.