STUDIES

The spatial spillover effects of small and medium-sized enterprise clusters and industrial estates on the unemployment rate: evidence from Java

Dwi Agus Prastiwi
Central Bureau Statistics,
Jakarta, Indonesia

Khoirunurrofik Khoirunurrofik
(corresponding author)
Department of Economics,
Faculty of Economics and Business,
Research Cluster on
Energy Modeling and
Regional Economic Analysis
(RCEMREA),
Universitas Indonesia;
Institute for Economic and
Social Research,
Universitas Indonesia,
Jakarta, Indonesia
Email: khoirunurrofik@ui.ac.id

Reflecting the dominance of Java's manufacturing industry at the national level, this study examines the relationships between industrial agglomeration at the micro, small, and medium enterprise (MSME) cluster level, medium and large industrial estate (IE) industries, and the unemployment rate (UR). The study investigates the spread and backwash effects that arise from industrial agglomeration and resulting spatial externalities on surrounding areas. The panel dataset covering 119 districts/cities in Java, Indonesia from 2014 to 2021 includes degrees of agglomeration in IE areas and among MSMEs, regions' unemployment rate, and regional socioeconomic variables. The study applies exploratory spatial data analysis (ESDA) and a spatial autoregressive (SAR) model to describe and examine spatial externalities. The ESDA reveals highly centralized IEs and MSMEs in northern and central Java, respectively. The SAR model demonstrates that IE agglomeration has a significantly positive direct effect on the unemployment rate, whereas MSME agglomeration has a negative but statistically negligible impact. The positive correlation between agglomeration in industrial areas and the unemployment rate indicates that the quality of current human resource mechanisms does not match the industry expectations, and the negative correlation between MSME agglomeration and unemployment rate reinforces this challenge.

Regional Statistics, Vol. 15. No. 2. 2025
Online first: Prastiwi–Khoirunurrofik 1–21; DOI: 10.15196/RS150201

These insights suggest a need to develop integrated policies to address regional unemployment challenges considering the decentralized authority of each district/city. The significant spatial externalities of unemployment rates between regions and the repercussions of industrial areas' agglomeration in Java necessitate effective labour market integration and regional cooperation. The study demonstrates this need by examining the different labour absorption mechanisms between IEs and MSMEs for confronting unemployment challenges.

Keywords:
spatial regression,
agglomeration,
industrial estate,
MSMEs,
unemployment

Online first publication date: 21 February 2025

Introduction

The unemployment rate is unique among the three macroeconomic indicators, with the most direct and immediate impact at the micro level (Mankiw 2016). According to 2021 Statistics Indonesia by Central Bureau Statistics (BPS), Indonesia's national open unemployment rate remains relatively high at 6.49%. Additionally, the disparities between regions remain high, with differences between the lowest and highest unemployment rates that are more than threefold. Significant unemployment rate differences among Indonesian provinces have prevailed for many years. In 2011, the gap between the lowest and highest unemployment rates increased nearly fivefold. Although the difference has decreased, reductions have been insignificant and a recurring pattern in regions with the lowest and highest unemployment rates has exacerbated inequality among Indonesian provinces. Therefore, it is essential to conduct a regional analysis across administrative boundaries to advance the development of scientifically informed macroeconomic policies to overcome the nation's unemployment problems (Formánek–Hušek 2016). Improving the balance of labour market supply and demand can solve these challenges. At the macro level, demand improvement is found to be a more effective and sustainable approach for minimizing unemployment (Dosi et al. 2019). Based on its proportion in the nation's GDP, the industrial sector has become the main driver of national economic development owing to the most significant GDP formation, with 20.55% in 2021.

Rapid industrial growth is expected to create higher demand, which diminishes the unemployment rate (Borjas 2016). However, advancing industrial development in all regions is implausible because of cost minimization and industrial development regulations. This fosters the tendency toward clustered industrial development and industrial agglomeration areas. 2021 Village Potential (Podes) data, published by

Central Bureau Statistics (BPS), indicate that Indonesia's micro, small, and medium enterprises (MSMEs) are highly concentrated in Java (55.13%). Industrial estates (IEs) are also concentrated in Java (23.19%) and Sumatra (46.13%). In addition to being the center of Indonesia's industry, Java has a reliable transportation and communication infrastructure that outperforms other Indonesian regions (BPS 2021). The superior transportation and communication infrastructure in Java should facilitate input–output flow across economic activities to establish equitable economic development among the regions (Kolomak 2020).

Forming an economic center can foster spread and backwash effects to surrounding areas (Chiang 2018). Regions that offer complementary goods or services for neighbouring regions can have a positive (spread) effect. Conversely, regions that offer substitute goods or services to other regions produce a negative (backwash) effect (Hirschman 1958). Previous spread-backwash effect-related studies have not focused on the relationship between unemployment rate and industrial agglomeration in a particular area. Some studies have examined government-defined industrial areas, while others have applied urban or metropolitan definitions as economic centers (Chiang 2018, Kamei–Nakamura 2022, Ke–Feser 2010, Tian et al. 2021).

As industrial agglomeration is not only government-driven but may also occur naturally and in industrial and non-industrial sectors that emerge in urban and metropolitan areas, the number of industries included in this study fills the gap in the existing literature on this research area. In addition, the majority of previous studies have concentrated on a single industry type such as MSMEs (Sarmah et al. 2021) or large industries (De Catris–Pallegrini 2015, Kwasny et al. 2019). Therefore, we employ an industrial number variable that distinguishes the scale of industries comprising IEs and MSMEs as predictors of the causes of regions' UR.

This study investigates the impact of IEs and MSME agglomeration on Java's unemployment rate, contributing to the growing body of research examining the differing mechanisms of labour absorption between the two types of industries for addressing unemployment challenges in Indonesia, particularly in Java. This study also identifies the locations of IE and MSME agglomerations in Java, which can support evidence-based infrastructure development to ensure that existing industrial agglomeration can produce positive externalities in its own and neighbouring regions.

The remainder of this paper is organized as follows. First, the conceptual frameworks and empirical studies that have applied variables relevant to this study are reviewed. Then, the authors outline the study's research methodology (data description, data sources, and method of analysis). Next, results and subsequent discussion are presented. Finally, the implications for Indonesian government policies related to industrial agglomeration and the unemployment rate are concluded and examined.

Literature review

The manufacturing industry is one of Indonesia's most labour-intensive sectors. Unfortunately, due to uneven development, some areas have fewer established industries than others, with a more significant number representing "industrial agglomeration areas" (O'Donoghue–Gleave 2004). Industrial agglomerations result from long-term trajectories that begin with the establishment of an industry in an area. With positive externalities, the location continues to expand and eventually forms an increasingly sizable industrial cluster (Phelps 1992).

Industrial agglomeration researchers have frequently referenced two widely adopted theories, the Marshall–Arrow–Romer (MAR) and Jacobs' externalities that focus on the externalities of industrial agglomeration. The MAR theory holds that positive externalities result from similar industries in a particular geographic area (specialization). In contrast, Jacobs' theory posits that positive externalities result from a wide range of industries in a specific area (diversity) (Glaeser et al. 1992, Henderson 1997). Each industry has specific externalities. MSME's access to the goods and services market tends to be easy because of lower *sunk cost*, which is often found in start-up industries (Khoirunurrofik 2020). Start-up industries tend to grow in heterogeneous environments (Duranton–Puga 2001) because they primarily focus on producing new products and developing market innovations (Hartog et al. 2012). Therefore, the externalities emerging from MSMEs represent Jacobs' model (Henderson 1997, Khoirunurrofik 2018). Conversely, large companies in industrial areas have a higher *sunk cost* as they are more established and classified as older or senior companies (Khoirunurrofik 2020). This type of industry tends to be localized and grow in a homogenous environment (Duranton–Puga 2001). This occurs because advanced industries typically focus on more efficient and cheaper product innovations (Hartog et al. 2012); hence, the externalities generated in IEs correspond with the MAR model (Henderson 1997, Khoirunurrofik 2018).

One of the externalities associated with industrial agglomeration is inter-firm knowledge spillover facilitated by the collection of industries in one location that can stimulate greater innovation (Beaudry–Schiffauerova 2009, Henderson 1997, Porter 2000). Additionally, industries in agglomeration areas often initiate intra-industry collaborations, which reduces input and output transportation costs and allows firms to benefit from a more specialized and efficient labour market (Henderson 1997). With increasing innovation and decreasing production costs, companies may expand their production scales (Borjas 2016, Grossman–Helpman 1991, Romer 1990).

The higher a company's production scale is, the more workers it requires (Prachowny 1993). High labour demand in a region does not necessarily correspond with local labour supply, which causes labour mobility. The theory of labour market imbalance asserts that labour migration mobility has a substantial effect on the convergence of labour supply between regions. Furthermore, commuters can balance labour demand and supply by widening the scope of the labour market itself (Niebuhr et al. 2012). The contemporary development of interregional transportation infrastructure and the relatively high cost of living in economic centers, workers' commuting costs have decreased (Sato 2000). This has contributed to a shift in worker mobility patterns from migration, resulting in a backwash effect of agglomeration to commuting, with a spread effect from industrial agglomeration (Burda–Hunt 2001, Rouwendal 1998).

Research method

We selected 119 districts/cities in Java for the 2014–2021 period, applying a spatial analysis technique to examine industrial agglomeration that requires an IE with an area of more than 20 ha. The dependent variables in this study are the unemployment rate and the proportion of the unemployed labour force to the total labour force, which are obtained from the National Labour Force Survey (SAKERNAS). The primary independent variable is industrial agglomeration, where the most frequently used measure has been the location quotient indicator (O'Donoghue–Gleave 2004). Therefore, the industrial agglomeration variable applied in this study is the proportion of IEs and MSMEs in a district or city in Java, as shown in the following equation:

$$aglo_i = \frac{ind_i}{ind_T} \times 100\% \qquad (1)$$

where $aglo_i$ is the magnitude of industrial agglomeration in district or city i, ind_i is the number of IEs or MSMEs in district or city i, and ind_T is the total number of IEs or MSMEs in Java. Data on the number of IEs and MSMEs are obtained from the triennial Podes, and the previous years' Podes data are used in years with no Podes data.

In addition to the main independent variables, several independent control variables are used to reduce bias in the parameter estimation. These variables include the gross regional domestic product (GRDP), percentage of industry to GRDP, consumer price index (CPI), and the percentage of urban areas that represent labour demand. Population density, average years of schooling, and labour force participation rate (LFPR) are also used to represent labour supply and the regency/city minimum wage. Table 1 presents the detailed definitions and data sources for the study's variables.

Table 1

Operational definitions of research variables

No	Variables – labels	Operational definitions	Units	Sources
		Dependent variables		
1	Unemployment rate – UR	$\frac{unemployment}{workforce} \times 100$	Percent	Region in figures
		Main independent variables		
2	Industrial estate – IE	Industrial activity areas covering at least 20 ha that are equipped with supporting facilities and infrastructure and developed and managed by an industrial estate company with an industrial estate business license in a district/city.	Total	District Podes 2014, 2018, and 2021
3	Micro, small, and medium enterprises – MSME	Industries with a workforce of less than 20 people in a district/city.	Total	Village Podes 2014, 2018, and 2021
		Independent control variables		
4	Gross regional domestic product – GRDP	The sum of all components of gross value added generated by economic sectors from various production activities. Goods and services are calculated using constant 2010 prices.	Billion rupiah	Region in figures
5	Percentage of industry to GRDP – PINDGRDP	$\frac{GRDP\ of\ manufacturing\ industry}{total\ GRDP} \times 100$	Percent	Web statistics Indonesia
6	Consumer price index – CPI	The average price from a collection of prices of goods and services consumed by residents/households in a certain period. The types of goods and services include foodstuffs; prepared food, beverages, and tobacco products; housing, water, electricity, gas, and fuel; clothing; health; education, recreation, and sports; transportation, communication; and financial services.	Point	Web statistics Indonesia
7	Region classification – URB	Proportion of villages with urban status in a district/city.	Percent	District Podes 2014, 2018, and 2021
8	Population density – POP	$\frac{population}{area}$	Per km^2	Region in figures
9	Average years of schooling – EDUC	Number of formal years of schooling completed by the population aged 15 years and over (excluding repeated years).	Year	IPM publications
10	Labour force participation rate – LFPR	$\frac{workforce}{population\ 15+} \times 100$	Percent	Region in figures
11	District/city minimum wage – WAGE	Employers using the minimum standard to pay employees or labourers in their business or work environment.	Rupiah	Wagepedia Ministry of Manpower and Governor Decree

We employ exploratory spatial data analysis (ESDA) to identify industrial agglomeration areas and clusters of open unemployment in Java. To do so, this study implements spatial autocorrelation to examine whether the relationship between location similarity in spatial proximity matches the similarity of observed variable values using ESDA. The statistical indicators used in the ESDA include the global Moran's index (Moran's I) and local indicators for spatial association (LISA) (Anselin 1999). The ESDA analysis is presented using a LISA quadrant distribution map to illustrate the number of IEs and MSMEs in 2021. We also present a bivariate LISA quadrant distribution map between unemployment rate, MSME, and IEs between 2014 and 2021.

In addition to the LISA quadrant distribution map, we calculate Moran's I and test for significance to reveal the magnitude of spatial autocorrelation according to the variables in the study. The Moran's I has a value of −1 to 1, where an index greater than 0 indicates positive spatial autocorrelation and a data pattern that forms a cluster. Conversely, a Moran's I that is less than 0 indicates a negative spatial autocorrelation and a scattered data pattern. A Moran's I of 0 indicates no spatial autocorrelation. The Moran's I calculation is as follows:

$$I = \frac{N \sum_i^N \sum_j^N W_{ij}(Y_i - \bar{Y})(Y_j - \bar{Y})}{\sum_i^N \sum_j^N W_{ij}(Y_i - \bar{Y})^2} \tag{2}$$

where $\bar{Y}$ is the average number of observations across regions, and W_{ij} is the spatial weight between regions *i* and *j*.

We use panel data spatial regression analysis to estimate the effects of independent variables on the dependent variable. Parameter estimation in panel data spatial regression follows almost the same rules as those in cross-sectional data. The only difference is the addition of time identity to all variables and errors in the model. As in panel data analysis, this study treats the spatial and time effects in this model as fixed or random effects, and model selection between fixed and random effects uses the Hausman test (Elhorst 2014).

Spatial regression emerged from empirical development of classical linear regression, introducing spatial considerations to parameter estimation calculations. The addition of spatial properties in spatial regression is based on the first law of geography presented by Tobler (1970) in which everything is always related, but things that are in closer proximity will have a stronger connection than those that are farther from each other. The initial models developed for spatial regression analysis examined spatial interactions with the dependent variable, which is usually referred to as the spatial autoregressive (SAR) model and spatial interactions with errors, which is called the spatial error model (SEM) (Griffith–Anselin 1989). The development of these models revealed three types of spatial interactions that may occur in spatial econometric models, which include endogenous interactions (spatial interactions that occur with the dependent variable), exogenous interactions (spatial interactions that occur with independent variables), and spatial interactions that occur in errors.

The spatial regression model that combines these three spatial interactions is referred to as the spatial Durbin model (SDM) (Elhorst 2014).

The first spatial regression model developed in this study is the SDM. This model has the most complete spatial interactions compared with the SAR model and the SEM. However, we still compare the SDM used in this study with non-spatial and other spatial models comprising SAR and SEM. The SDM equation used in this study was as follows:

$$Y = \rho WY + \alpha 1_N + X\beta + WX\theta + \varepsilon \quad (3)$$

where Y is a vector of dependent variables, which is the unemployment rate of all districts/cities in Java in this study; $a1_N$ is a vector of constant parameters (a) to be estimated; X is a matrix of independent variables; β is a vector of associated parameters; W is the spatial weight matrix; ρ is the spatial autoregressive parameter for the endogenous interaction effect (WY) that captures the spatial interaction associated with the dependent variable; θ is the spatial parameter vector for the exogenous interaction effect (WX); and ε is the error with zero mean and constant variance.

We employ a spatial weight matrix based on inverse distance to define neighbouring areas in the ESDA descriptive analysis and spatial regression. This technique is chosen because the distance spatial weighting matrix overcomes Java's topological problems in which several districts/cities are located on separate islands (De Catris–Pallegrini 2015, Getis 2009). Furthermore, because this study's focus on the UR, which is closely related to labour mobility, using the distance between regions to define neighbouring relationships among regions is considered more appropriate than using direct boundaries between regions (Jeguirim 2021).

Results and discussion

Descriptive analysis

The average unemployment rate of districts/cities on Java from 2014 to 2021 is still relatively high at 6%. Although most districts/cities in Java do not have IEs, some host a large proportion. This trend is demonstrated by the average number of IEs at 0.6, minimum value of 0, and maximum of 29. Slightly different from IEs, although disparities in the number of MSMEs remain, the 2014–2021 data suggest that every district/city in Java had MSMEs, which is indicated by the minimum value of 6, the maximum of 50,911, and the average of 9,563. Although a considerable range between the minimum and maximum values is evident, the standard deviation (8,476.27), which is still less than the average value, indicates a slight variation in the MSME data. Furthermore, the independent control variables used in this study generally present subtle data variations, as shown by standard deviations that are smaller than mean values (Table 2).

Table 2

Descriptive statistics of variables in the study (Java, 2014–2021)

Variables	Mean	Standard deviation	Minimum	Maximum
UR	6.00	2.68	0.85	14.87
IND	0.60	2.20	0.00	29.00
MSME	9,563.45	8,476.27	6.00	50,911.00
WAGE	2,087,770.00	863,106.60	910,000.00	4,798,312.00
GRDP	49,941.27	76,771.42	2,624.24	460,081.00
PINDGRDP	23.56	17.85	0.84	81.09
CPI	128.84	9.36	110.04	154.08
POP	3,235.54	4,372.58	277.00	20,813.00
EDUC	8.07	1.66	3.49	11.82
URB	50.48	33.81	3.47	100.00
LFPR	67.44	4.46	53.77	80.64

Spatial cluster identification

The quadrants of each region are represented by five different colours to indicate the different types of areas on the LISA map. Areas in white are too statistically insignificant to fall into one of the four quadrants of the LISA. Dark red indicates an area of high concentration, meaning that the area has a high variable value and is surrounded by neighbouring areas with the same value. Dark blue is an area of low concentration, which means that the area has a low variable value and is surrounded by neighbouring areas with low variable values. The pink region is a high-outlier region, meaning that the region has a high variable value but is surrounded by neighbouring regions with low variable values. Light purple indicates a low-outlier region, meaning that the region has a low variable value but is surrounded by neighbouring regions with high variable values.

Figure 1 illustrates the distributions of areas according to the number of IEs in 2021, which was dominated by concentrated areas (15.97%) compared to outlier areas. Areas with high concentrations were more widely distributed in the western region of Java (West Java and Banten). The possible rationale for this finding is that IEs must be developed in strategic locations, including nearby cities, while considering the convenience of community settlements. The first IE development in Indonesia, the Jakarta Industrial Estate Pulo Gadung, was established in 1973 in the western part of Java.

Regarding the number of IEs, low-concentration areas are more widely distributed in the central and eastern regions of Java (Central Java, DI Yogyakarta, and East Java). Furthermore, as shown in Figure 2, regional MSME distribution in 2021 was dominated by concentrated rather than outlier areas. In contrast to IEs' regional distribution pattern, high-concentration MSME areas were more widely spread in the Central Java Province. Meanwhile, low-concentration areas were more widely distributed in the western region of Java (including Jakarta, West Java, and Banten).

Figure 1
Local index of spatial association (LISA) based on the number of IEs (Java, 2021)

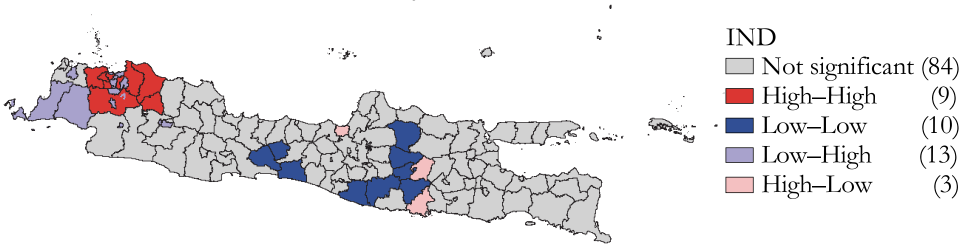


Figure 2
Local index of spatial association (LISA) based on the number of MSMEs (Java, 2021)

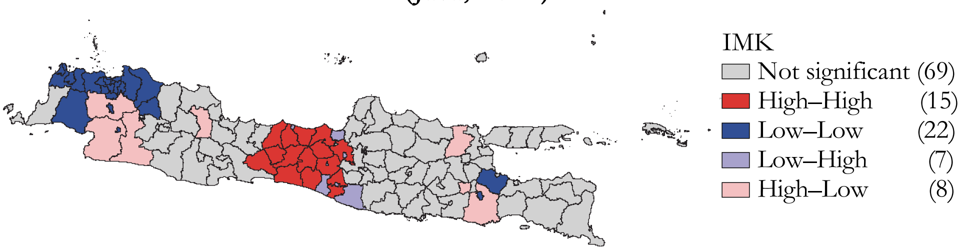


Figure 3
Local index of spatial association (LISA) based on unemployment rate and the number of IEs (Java, 2014)

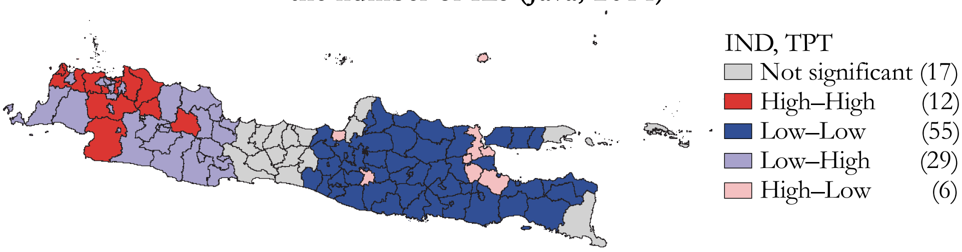


Figure 4
Local index of spatial association (LISA) based on unemployment rate and the number of IEs (Java, 2021)

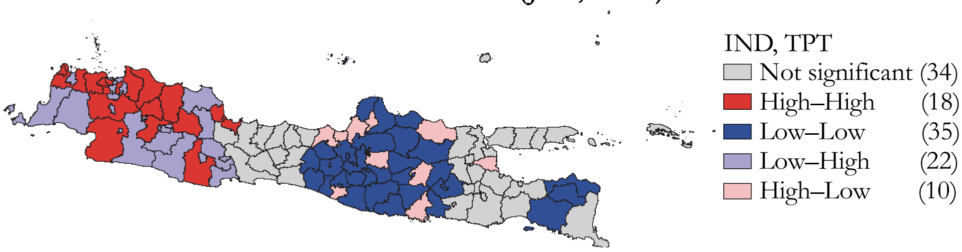


Figure 5

Local index of spatial association (LISA) by unemployment rate and the number of MSMEs (Java, 2014)

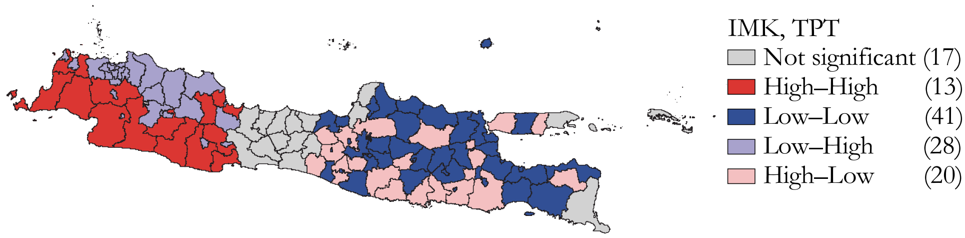
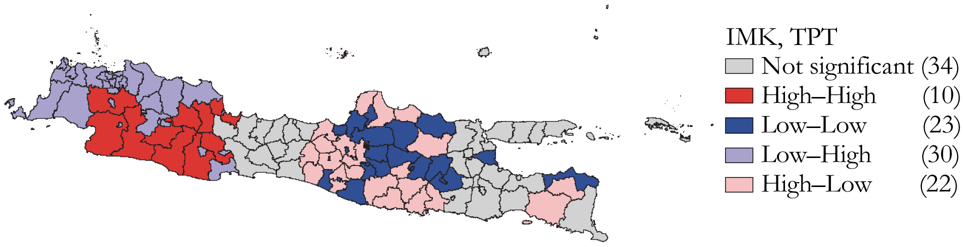


Figure 6

Local index of spatial association (LISA) based on unemployment rate and the number of MSMEs (Java, 2021)



Source: ESDA results (2023).

Figures 3 and 4 show the distribution patterns of LISA clusters for IEs, revealing that concentration and deficient concentration areas dominated the unemployment rate in 2014 and 2021. This trend suggests that most districts/cities had few IEs and are surrounded by neighbouring areas with low UR. These areas are primarily located in the central and eastern parts of Java Island, which are low-concentration areas based on the number of IEs. The dominance of concentration areas in 2014 and 2021 indicates a positive spatial autocorrelation between IEs and UR. This demonstrates that areas with many IEs tend to be surrounded by areas with high URs and vice versa. In other words, the spillover of IE on an area's unemployment rate is likely to represent a backwash effect.

In contrast to previous findings, a change is evident in the pattern of LISA cluster distribution according to MSMEs and unemployment rate in 2014 and 2021. In 2014, the distribution of MSME clusters according and URs in Java was dominated by centralized areas (45.38%), indicating a positive spatial autocorrelation between MSME and unemployment rate (Figure 5). However, as shown in Figure 6, blue and pink areas dominate, indicating that most districts/cities in Java were outlier areas in 2021. This finding reveals that the spatial autocorrelation relationship between MSME and unemployment rate has become negative, showing that areas with a high number of MSMEs tended to be surrounded by areas with low URs, and vice versa in 2021. In other words, the MSME spillover on an area's unemployment rate in 2021 is likely to represent a spread effect.

Econometric spatial analysis

Table 3 details the results of the significance test and Moran's I values for all variables used in this study. Table 3 also illustrates the positive significant Moran's I value of unemployment rate for all years, revealing that the Moran's I value of unemployment rate in Java increased from 2014 to 2019, indicating that the geographical distribution of unemployment in Java has become increasingly centralized over time. However, an opposite pattern occurred from 2020 to 2021, as the value of Moran's I for unemployment rate decreased, indicating that the geographical distribution of unemployment in Java in that period was increasingly spread out. This intriguing result can be attributed to the Covid-19 pandemic between 2020 and 2021, which resulted in extended lockdowns, which were one of the most substantial causes of the increase in regional unemployment rates (Houston 2020). Therefore, regions with previously low URs also witnessed increased URs, which elevated dispersion between regions.

Table 3 presents the significant positive effects of Moran's I on the IE and MSME variables. However, from 2018 to 2020, the Moran's I value for the IE variable was not statistically significant. The table also presents the minimum Moran's I values for these two variables. This finding could be attributed to the identification of industrial agglomerations that typically had small coverage areas (De Catris–Pallegrini 2015). Accordingly, industry agglomeration should be identified using individual- or firm-level data (Guimarães et al. 2011, Marcon–Puech 2017); however, when facing limitations of spatial data at the trim level, aggregate data can be used as a proxy to illustrate the spatial relationships between variables (De Catris–Pallegrini 2015). Furthermore, Table 3 shows that all independent control variables in this study, except for the percentage of industry in the GRDP, have positive and significant Moran's I values.

Table 3

Moran's I by variables

Variables	2014	2015	2016	2017	2018	2019	2020	2021
UR	0.49***	0.54***	0.63***	0.66***	0.62***	0.68***	0.58***	0.50***
IND	0.07**	0.07**	0.07**	0.07**	0.02	0.02	0.02	0.06**
MSME	0.17***	0.17***	0.17***	0.17***	0.18***	0.18***	0.18***	0.13***
WAGE	0.72***	0.69***	0.69***	0.68***	0.68***	0.68***	0.68***	0.68***
POP	0.20***	0.20***	0.20***	0.21***	0.21***	0.21***	0.20***	0.20***
EDUC	0.16***	0.17***	0.18***	0.18***	0.18***	0.19***	0.20***	0.18***
URB	0.11***	0.10***	0.11***	0.11***	0.11***	0.11***	0.11***	0.11***
LFPR	0.41***	0.48***	0.51***	0.43***	0.52***	0.46***	0.42***	0.5337***
GRDP	0.22***	0.21***	0.21***	0.21***	0.21***	0.21***	0.21***	0.2140***
PINDGRDP	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00
CPI	0.25***	0.40***	0.44***	0.45***	0.49***	0.57***	0.61***	0.5942***

Note: statistical significance at 1% (***), 5% (**), and 10% (*).

Before conducting model comparisons, Table 4 presents a comparison of the results of data processing using several different spatial weighing matrices, including a spatial weighing matrix based on distance with a threshold distance of 0.1–10 km, a queen contiguity spatial weighing matrix, and an inverse distance spatial weighing matrix. Table 4 reveals that Akaike information criterion (AIC), Bayesian information criterion (BIC), and R-squared values between all spatial weighting matrices do not differ widely. In addition, the magnitude and sign of the parameter estimates for the main independent variables are also quite similar. Considering that district/city areas located on different islands can still be included in the analysis and the theoretical assumptions noted in the research methodology, this study uses the inverse distance spatial weight matrix.

Table 4

Comparison of spatial weighting matrices

Spatial weighting matrices		R^2 between	βIND	βMSME	ρW	Number significant variables (α = 0.1)	AIC	BIC
Distance to a certain threshold	0.1	0.4705	0.0328	0.0812	0.3784***	4	3,306.099	3,422.705
	0.2	0.4635	0.0294	0.0319	0.3888***	4	3,228.852	3,345.457
	0.3	0.4917	0.0367*	0.0443	0.4366***	5	3,170.631	3,287.236
	0.4	0.5572	0.0360*	−0.0425	0.5282***	5	3,098.186	3,214.791
	0.5	0.6012	0.0317*	−0.0397	0.6026***	4	3,009.786	3,126.391
	0.6	0.6276	0.0343*	−0.0908	0.6596***	4	2,965.106	3,081.712
	0.7	0.6279	0.0352*	−0.0856	0.6897***	5	2,954.065	3,070.671
	0.8	0.6391	0.0346*	−0.0759	0.7132***	4	2,942.844	3,059.45
	0.9	0.6325	0.0349*	−0.0883	0.7264***	6	2,939.328	3,055.934
	1	0.6504	0.0368**	−0.0921	0.7352***	4	2,938.86	3,055.466
	2	0.762	0.0363**	−0.101	0.7879***	4	2,910.716	3,027.321
	3	0.7839	0.0352*	−0.1183	0.8177***	5	2,902.941	3,019.546
	4	0.7847	0.0342*	−0.1189	0.8298***	5	2,900.615	3,017.221
	5	0.7715	0.0325*	−0.0941	0.8392***	5	2,906.629	3,023.235
	6	0.7587	0.0311*	−0.0886	0.8441***	5	2,908.525	3,025.131
	7	0.7469	0.0311*	−0.0830	0.8454***	5	2,912.43	3,029.036
	8	0.7404	0.0309*	−0.0828	0.8448***	5	2,914.716	3,031.322
	9	0.7388	0.0308*	−0.0843	0.8447***	5	2,915.533	3,032.139
	10	0.7388	0.0308*	−0.0843	0.8447***	5	2,915.533	3,032.139
QUEEN		0.6268	0.0294	−0.0938	0.5665***	2	3,061.205	3,177.811
IDISTANCE		0.7388	0.0308*	−0.0844	0.8447***	5	2,915.533	3,032.139

Note: statistical significance at 1% (***), 5% (**), and 10% (*).

We next compare the non-spatial ordinary least squares (OLS) model with the spatial model when spatial autocorrelation is identified in the variables. The comparison of the AIC and BIC values can be used to identify the best model. A smaller AIC/BIC value indicates a better model (Anderson–Burnham 2002). Table 5 reveals smaller AIC and BIC values for the spatial model. Furthermore, the SAR model had the lowest AIC and BIC values compared with other spatial models.

Having analyzed AIC and BIC values, a further test is conducted on each spatial weighting parameter to determine which of the three spatial models is most suitable for the study. If $\theta = 0$, the SAR model is considered better than SDM, and if $\theta = -\delta\beta$, the SEM model is considered better than SDM (Belotti et al. 2017, Elhorst 2014). These tests reveal chi-squared values for the comparison of SDM and SEM of 26.83 (0.0028), indicating that the SDM is more accurate than SEM. The comparison of SDM and SAR demonstrates chi-squared values of 13.31 (0.2070), indicating that the SAR model is more accurate than the SDM. Therefore, based on the comparison of AIC and BIC values and spatial model comparisons, we conclude that the best model for this study is the SAR model.

Table 5

Estimation results of spatial and non-spatial models

Criteria		OLS	SAR	SEM	SDM
Main	IND	0.1166***	0.0328*	0.0298	0.0308*
	MSME	−0.0519	−0.0240	0.062	−0.0844
	WAGE	1.7182***	1.7051***	3.6629***	1.0138
	GRDP	−0.1911**	0.0277	0.0047	0.0796
	PINDGRDP	0.0141***	0.0116	0.0047	0.0133*
	CPI	0.0191**	−0.0308***	−0.0290	−0.0136
	LFPR	−0.3177***	−0.0699***	−0.0768***	−0.0687***
	POP	0.0001***	−0.0001	−0.0000	−0.0001
	EDUC	−0.1843**	−0.0656	−0.1607	−0.0666
	URB	0.0032	0.0139**	0.0149*	0.0137**
	_cons	2.7015	−15.8375***	−37.7957***	−31.1493*
Spatial rho			0.8749***		0.8447***
Spatial lambda				0.8697***	
Wx	IND				−0.1835
	MSME				1.5570*
	WAGE				2.0479
	GRDP				−0.5786
	PINDGRDP				0.0691
	CPI				−0.0485
	LFPR				−0.0468
	POP				−0.0004
	EDUC				0.2498
	URB				0.0928***
R^2 between		0.5489	0.6655	0.5251	0.7388
AIC		3,834.702	2,908.61	2,977.556	2,915.533
BIC		3,888.147	2,976.629	3,045.576	3,032.139
Log-likelihood			−1,440.3048	−1,474.7779	−1,433.7667

Note: standard errors are in parentheses. Statistical significance at 1% (***), 5% (**), and 10% (*).

Direct and indirect effects

The SAR analysis decomposes the total impact into the direct and indirect effects of each independent variable in the model. Table 6 shows the influence of all significant independent variables on the dependent variable in the same direction, indicating the spillover process. The results show that magnitude of the direct and indirect effects on each independent variable was greater than that of the coefficient of each independent variable in the model. The findings reveal the importance of including feedback effects in the model as it strengthens the effect of each exogenous variable on the unemployment rate (LeSage–Pace 2009).

Table 6

Direct and indirect effects and total of the SAR model

Variables	Direct	Indirect	Total
IND	0.0364*(0.0201)	0.2481(0.1592)	0.2845(0.1767)
MSME	−0.0320(0.1312)	−0.2335(0.9733)	−0.2656(1.0987)
WAGE	1.9000***(0.4948)	12.9030***(5.1779)	14.8030***(5.4783)
GRDP	0.0216(0.1692)	0.1725(1.2501)	0.1941(1.4120)
PINDGRDP	0.0126(0.008)	0.0867(0.0665)	0.0993(0.0735)
CPI	−0.0341***(0.0114)	−0.2313**(0.1050)	−0.2654**(0.1131)
LFPR	−0.0759***(0.0162)	−0.5231***(0.2140)	−0.5990***(0.2242)
POP	−0.0000(0.0000)	−0.0005(0.0004)	−0.0005(0.0004)
EDUC	−0.0678(0.1329)	−0.4878(1.0020)	−0.5556(1.1267)
URB	0.0149**(0.0069)	0.1041(0.0670)	0.1190*(0.0726)

Note: standard errors are in parentheses. Statistical significance at 1% (***), 5% (**), and 10% (*).

The data demonstrate that IE agglomeration has a more dominant positive impact on the UR, with an impact size of 0.0364, indicating that a 1% increase in IE agglomeration will increase the unemployment rate by 0.04%; however, the indirect impact was not statistically significant (Table 6). This result aligns with the findings of several studies demonstrating that the direct effect of industrial development is an increased number of unemployed people that is attributable to the development of automated machine technologies (e.g., Calvino–Virgillito 2018). Likewise, IE development has reduced agricultural land, necessitating the transition of agricultural labour to the manufacturing industry (Yang 2014). Unfortunately, more prominent industries require labour with higher skills that sometimes cannot be accommodated by local labour, which eventually results in migration flow (De Catris–Pallegrini 2015). Continuous migration to agglomeration areas causes such areas to be unable to absorb all incoming labour, which increases unemployment (Epifani–Gancia 2005).

This condition mirrors Statistics Indonesia's data demonstrating that West Java is a high-concentration area of IEs and the leading destination for lifetime and return migrants in Indonesia, as the results of our ESDA analysis also indicate. Migrants, who are predominantly people of productive age (20–39 years old) and have higher education than non-migrants (Statistics Indonesia 2021), are primarily labour migrants seeking employment in IE agglomeration areas (Statistics Indonesia 2023).

Table 6 shows that MSME agglomeration does not have any significant direct or indirect effect on the unemployment rate. This finding has been confirmed by previous research demonstrating that a company's age, rather than MSME status, influences labour growth (Haltiwanger et al. 2013). However, although insignificant, the MSME parameter exhibits a negative sign, suggesting that a higher proportion of MSMEs in a district/city lowers the unemployment rate. This finding may be related to the role of MSMEs in absorbing labour, which is extremely significant for developing countries, considering MSMEs' limited technology and labour-intensive nature (Sarmah et al. 2021). This finding aligns with the fact that agglomeration can facilitate the job–education matching process and reduce UR, in particular for vocational secondary school graduates (Paramitasari et al. 2024). However, MSMEs are an economic business field with more significant risks due to limited capital and technology; therefore, MSME actors also have a higher risk of becoming jobless.

This was especially evident in the period of economic turmoil between 2020 and 2021 due to Covid-19 pandemic, when the number of MSMEs in Indonesia decreased, as reported by Statistics Indonesia. Although large and medium industries also experienced a decrease, the decline in MSMEs in 2020 was higher (3.9%) compared with large and medium industries, which only decreased by 2.4%. Another possible explanation is that Indonesian MSMEs are vulnerable because most MSME workers (74.65%) are junior high school graduates or below. Entrepreneurs with a primary school education or below manage most MSMEs in Indonesia (52.90%) (Statistics Indonesia 2023). Low-qualified business actors illustrate limited capabilities and skills, which impedes MSME development.

Another finding is that wages and the proportion of urban areas positively influence the unemployment rate. Labour market equilibrium theory asserts that a minimum wage above the equilibrium wage raises unemployment (Chu et al. 2020). Furthermore, the perceived comfortable life in urban areas drives rural residents' to migration, resulting in more competitive job-seeking and increasing unemployment rates (Lyu et al. 2019).

The findings also indicate that the CPI and LFPR have significant direct and indirect negative influence on the unemployment rate. This result was confirmed by Philips' theory, which argued that a trade-off exists between inflation and the unemployment rate (Bhattarai 2016). Previous research has also revealed a negative relationship between unemployment and the LFPR because a lower fertility rate reduces the productive population composition (Feng et al. 2017). Increasing labour force competition has also been found to raise the number people leaving the labour force to pursue higher education (Liu 2012), which results in a labour force with low education and skills that are difficult to absorb into the labour market, raising the unemployment rate (Feng et al. 2017, Liu 2012).

Conclusions and implications

Most of the districts/cities on Java are low-concentration areas for IEs, which are primarily dispersed in Java's central and eastern regions. The geographical distribution of concentration areas and outliers for MSMEs reveals low-concentration areas primarily in the western and eastern regions of Java. Based on the calculation and Moran's I significance testing, we determine that almost all the variables used in this study had significantly positive spatial autocorrelation from 2014 to 2021. Based on the results using the SAR model, we find that the agglomeration of IEs had a significant positive impact on the UR, indicating that IE agglomeration increases the UR, particularly in regions where industrial agglomerations are located.

In contrast, MSME agglomeration exhibits direct and indirect negative impacts on unemployment; however, this impact is statistically insignificant. Other factors that significantly affect regions' unemployment rate are the minimum wage and proportion of urban areas, which have a positive impact, and the CPI and LFPR, which have significantly negative direct and indirect effects. The findings reveal that the agglomeration impact of industries of different economic sizes exert different effects on the provision of employment opportunities and efforts to reduce unemployment.

These findings are essential for developing integrated regional policies to overcome unemployment issues depending on each district/city's decentralized authority. Such action is pivotal because of the strong spatial relationship in the unemployment rate demonstrated among the regions. Another implication is that good market integration has been established because of the backwash effect caused by IEs' agglomeration in Java. Furthermore, market specialization increases interregional dependence and strengthens interregional market integration (Kolomak 2020). In addition, enhancing the workforce's skills is critical for remaining competitive and ensuring a labour force with the skills required by industries in the area.

As this research is limited to data aggregation for identifying industrial agglomerations, future research should focus on micro level analysis of enterprises or specified industries rather than aggregate data (Guimarães et al. 2011, Marcon–Puech 2017). Individual firms' spatial externalities, rather than regional externalities, can be measured using micro data. In addition, it is also essential to note that regions that are integrated with other regions in spatial analysis will have greater relationships than isolated regions (Rae 2009). Therefore, further research could differentiate the magnitude of externalities arising from industrial agglomeration between integrated and isolated regions.

Acknowledgement

The authors are grateful for the support of the PUTI Grant NKB-578/2023 of Universitas Indonesia.

REFERENCES

ANDERSON, D. R.–BURNHAM, K. P. (2002): Avoiding pitfalls when using information-theoretic methods *The Journal of wildlife management* 912–918.

ANSELIN, L. (1999): Interactive techniques and exploratory spatial data analysis. In: LONGLEY, P. A.–GOODCHILD, M. F.–MAGUIRE, D. J.–RHIND, D. W. (eds.): *Geographic information systems: principles, techniques, management and applications* pp. 253–266., Wiley.

BEAUDRY, C.–SCHIFFAUEROVA, A. (2009): Who's right, Marshall or Jacobs? The localization versus urbanization debate *Research Policy* 38 (2): 318–337. https://doi.org/10.1016/j.respol.2008.11.010

BELOTTI, F.–HUGHES, G.–PIANO MORTARI, A. (2017): Spatial panel data models using stata *SSRN Electronic Journal* 17 (1). https://doi.org/10.2139/ssrn.2754703

BHATTARAI, K. (2016): Unemployment-inflation trade-offs in OECD countries *Economic Modelling* 58: 93–103. https://doi.org/10.1016/j.econmod.2016.05.007

BORJAS, G. J. (2016): *Labor economics* McGraw-Hill Education, New York.

BADAN PUSAT STATISTIK [BPS] (2021): *Profil migran hasil survei sosial ekonomi nasional 2021* Jakarta, Indonesia.

BURDA, M. C.–HUNT, J. (2001): From reunification to economic integration: productivity and the labour market in Eastern Germany *Brookings Papers on Economic Activity* 2: 1–71. https://doi.org/10.1353/eca.2001.0016

CALVINO, F.–VIRGILLITO, M. E. (2018): The innovation-employment nexus: a critical survey of theory and empirics *Journal of Economic Surveys* 32 (1): 83–117. https://doi.org/10.1111/joes.12190

CHIANG, S. (2018): Assessing the merits of the urban-led policy in China: spread or backwash effect? *Sustainability* (Switzerland) 10 (2): 451. https://doi.org/10.3390/su10020451

CHU, A. C.–KOU, Z.–WANG, X. (2020): Dynamic effects of minimum wage on growth and innovation in a Schumpeterian economy *Economics Letters* 188: 108943. https://doi.org/10.1016/j.econlet.2020.108943

DE CATRIS, M.–PALLEGRINI, G. (2015): Agglomeration effects on regional unemployment in Europe *CREI Working Paper* 7. https://doi.org/10.13140/RG.2.1.3720.7125

DOSI, G.–PEREIRA, M. C.–ROVENTINI, A.–VIRGILLITO, M. E. (2019): What if supply-side policies are not enough? The perverse interaction of flexibility and austerity *Journal of Economic Behavior and Organization* 162: 360–388. https://doi.org/10.1016/j.jebo.2018.11.026

DURANTON, G.–PUGA, D. (2001): Nursery cities: urban diversity, process innovation, and the life cycle of products *American Economic Review* 91 (5): 1454–1477. https://doi.org/10.1257/aer.91.5.1454

ELHORST, J. P. (2014): *Spatial econometrics: from cross-sectional data to spatial panels* Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40340-8

EPIFANI, P.–GANCIA, G. A. (2005): Trade, migration and regional unemployment *Regional Science and Urban Economics* 35 (6), 625–644.

FENG, S.–HU, Y.–MOFFITT, R. (2017): Long run trends in unemployment and labor force participation in urban China *Journal of Comparative Economics* 45 (2): 304–324. https://doi.org/10.1016/j.jce.2017.02.004

FORMÁNEK, T.–HUŠEK, R. (2016): Spatial aspects of unemployment in the Visegrad-group economies *Creative and Knowledge Society* 6 (2): 1–12. https://doi.org/10.1515/cks-2016-0007

GETIS, A. (2009): Spatial weights matrices *Geographical Analysis* 41 (4): 404–410. https://doi.org/10.1111/j.1538-4632.2009.00768.x

GLAESER, E. L.–KALLAL, H. D.–SCHEINKMAN, J. A.–SHLEIFER, A. (1992): Growth in cities *Journal of Political Economy* 100 (6): 1126–1152. https://doi.org/10.1086/261856

GRIFFITH, D. A.–ANSELIN, L. (1989): Spatial econometrics: methods and models *Economic Geography* 65 (2): 160–162. https://doi.org/10.2307/143780

GROSSMAN, G.–HELPMAN, E. (1991): *Innovation and growth in the global economy* MIT Press, Cambridge, MA.

GUIMARÃES, P.–FIGUEIREDO, O.–WOODWARD, D. (2011): Accounting for neighboring effects in measures of spatial concentration *Journal of Regional Science* 51 (4): 678–693. https://doi.org/10.1111/j.1467-9787.2011.00723.x

HALTIWANGER, J.–JARMIN, R. S.–MIRANDA, J. (2013): Who creates jobs? Small versus large versus young *Review of Economics and Statistics* 95 (2): 347–361. https://doi.org/10.1162/REST_a_00288

HARTOG, M.–BOSCHMA, R.–SOTARAUTA, M. (2012): The impact of related variety on regional employment growth in Finland 1993–2006: high-tech versus medium/low-tech *Industry and Innovation* 19 (6): 459–476. https://doi.org/10.1080/13662716.2012.718874

HENDERSON, V. (1997): Externalities and industrial development *Journal of Urban Economics* 42 (3): 449–470. https://doi.org/10.1006/juec.1997.2036

HIRSCHMAN, A. O. (1958): *The strategy of economic development* Yale University Press, New Haven, CT. https://doi.org/10.2307/1235188

HOUSTON, D. (2020): Local resistance to rising unemployment in the context of the Covid-19 mitigation policies across Great Britain *Regional Science Policy and Practice* 12(6): 1189–1209. https://doi.org/10.1111/rsp3.12364

JEGUIRIM, K. (2021): An inquiry into the regional unemployment disparities in Tunisia: a spatial Durbin model *Middle East Development Journal* 13 (2): 265–291. https://doi.org/10.1080/17938120.2021.1958559

KAMEI, A.–NAKAMURA, S. (2022): Urban agglomerations and wage and self-employment jobs in Ethiopia *Journal of Development Studies* 58 (9): 1693–1710. https://doi.org/10.1080/00220388.2022.2043282

KE, S.–FESER, E. (2010): Count on the growth pole strategy for regional economic growth? Spread-backwash effects in greater central China *Regional Studies* 44 (9): 1131–1147. https://doi.org/10.1080/00343400903373601

KHOIRUNURROFIK, K. (2018): Local economic structure, productivity growth, and industry life cycle: evidence from Indonesia *Asia-Pacific Journal of Regional Science* 2 (2): 453–475. https://doi.org/10.1007/s41685-018-0072-6

KHOIRUNURROFIK, K. (2020): Does the crisis change the nature of agglomeration economies in Indonesia? A productivity analysis of pre-post 1997–1998 financial crisis *Region* 7 (2): 85–106. https://doi.org/10.18335/region.v7i2.308

KOLOMAK, E. A. (2020): Estimation of the spatial connectivity of the economic activity of Russian regions *Regional Research of Russia* 10 (3): 301–307. https://doi.org/10.1134/S2079970520030089

KWASNY, J.–MROCZEK, A.–ULBRYCH, M. (2019): Industrial clustering and economic performance. In search for evidence from Poland *Ekonomicko-Manazerske Spektrum* 13 (1): 109–119. https://doi.org/10.26552/ems.2019.1.109-119

LESAGE, J.–PACE, R. K. (2009): *Introduction to spatial econometrics* Chapman and Hall/CRC, New York. https://doi.org/10.1201/9781420064254

LIU, Q. (2012): Unemployment and labor force participation in urban China *China Economic Review* 23 (1): 18–33. https://doi.org/10.1016/j.chieco.2011.07.008

LYU, H.–DONG, Z.–ROOBAVANNAN, M.–KANDASAMY, J.–PANDE, S. (2019): Rural unemployment pushes migrants to urban areas in Jiangsu Province, China *Palgrave Communications* 5: 92. https://doi.org/10.1057/s41599-019-0302-1

MANKIW, N. G. (2016): *Macroeconomics* Worth Publishers, Broadway, UK.

MARCON, E.–PUECH, F. (2017): A typology of distance-based measures of spatial concentration *Regional Science and Urban Economics* 62: 56–67. https://doi.org/10.1016/j.regsciurbeco.2016.10.004

NIEBUHR, A.–GRANATO, N.–HAAS, A.–HAMANN, S. (2012): Does labour mobility reduce disparities between regional labour markets in Germany? *Regional Studies* 46 (7): 841–858. https://doi.org/10.1080/00343404.2010.532118

O'DONOGHUE, D.–GLEAVE, B. (2004): A note on methods for measuring industrial agglomeration *Regional Studies* 38 (4): 419–427. https://doi.org/10.1080/03434002000213932

PARAMITASARI, N.–KHOIRUNURROFIK, K.–MAHI, B. R.–HARTONO, D. (2024): Charting vocational education: impact of agglomeration economies on job-education mismatch in Indonesia *Asia-Pacific Journal of Regional Science* 8: 461–491. https://doi.org/10.1007/s41685-024-00333-x

PHELPS, N. A. (1992): External economies, agglomeration and flexible accumulation *Transactions – Institute of British Geographers* 17 (1): 35–46. https://doi.org/10.2307/622635

PORTER, M. E. (2000): Location, competition, and economic development: local clusters in a global economy *Economic Development Quarterly* 14 (1): 15–34. https://doi.org/10.1177/089124240001400105

PRACHOWNY, M. F. J. (1993): Okun' s law: theoretical foundations and revised estimates *The Review of Economics and Statistics* 75 (2): 331–336. https://doi.org/https://doi.org/10.2307/2109440

RAE, A. (2009): Isolated entities or integrated neighbourhoods? An alternative view of the measurement of deprivation *Urban Studies* 46 (9): 1859–1878. https://doi.org/10.1177/0042098009106019

ROMER, P. M. (1990): Endogenous technological growth *NBER Working Paper Series* 3210, Cambridge, MA.

ROUWENDAL, J. (1998): Search theory, spatial labor markets, and commuting *Journal of Urban Economics* 43 (1): 1–22. https://doi.org/10.1006/juec.1996.2033

SARMAH, A.–SAIKIA, B.–TRIPATHI, D. (2021): Can unemployment be answered by micro small and medium enterprises? Evidences from Assam *Indian Growth and Development Review* 14 (2): 199–222. https://doi.org/10.1108/IGDR-09-2020-0140

SATO, Y. (2000): Search theory and the wage curve *Economics Letters* 66 (1): 93–98. https://doi.org/10.1016/S0165-1765(99)00174-3

STATISTICS INDONESIA (2021): Hasil Sensus Penduduk 2020 *Berita Resmi Statistik* 24 (7/1). https://www.bps.go.id/id/pressrelease/2021/01/21/1854/hasil-sensus-penduduk--sp2020--pada-september-2020-mencatat-jumlah-penduduk-sebesar-270-20-juta-jiwa-.html

STATISTICS INDONESIA (2023): *Profil Industri Mikro dan Kecil 2021*. https://www.bps.go.id/id/publication/2023/04/18/228f9c6242ef5a06b2dc740e/profilindustri-mikro-dan-kecil-2021.html

TIAN, M.–LI, T.–YE, X.–ZHAO, H.–MENG, X. (2021): The impact of high-speed rail on service industry agglomeration in peripheral cities *Transportation Research Part D: Transport and Environment* 93: 102745. https://doi.org/10.1016/j.trd.2021.102745

TOBLER, W. R. (1970): A computer movie simulating urban growth in the Detroit region *Economic geography* 46 (1): 234–240.

YANG, X. (2014): Labor market frictions, agglomeration, and regional unemployment disparities *Annals of Regional Science* 52 (2): 489–512. https://doi.org/10.1007/s00168-014-0596-2