387-JEBE-Similarity- --

by zulfa.eef@gmail.com 1

Submission date: 12-Sep-2025 12:25PM (UTC+0530)
Submission ID: 2748712488
File name: 387-JEEB-Similarity--.docx (100.23K)
Word count: 3548
Character count: 19659

ABSTRACT

The informal employment sector remains a dominant part of the labor market in Indonesia, which is influenced by various socioeconomic factors. This study aims to analyze the spatial influence of poverty, education level, and internet access on the distribution of informal employments in 34 provinces in Indonesia using the Spatial Error Model (SEM). The results show significant spatial autocorrelation, especially for poverty and informal workers, which confirms the existence of inter-regional spillover effects. The SEM estimation shows that poverty has a positive and significant effect on the level of informal employment, while higher education levels and internet access significantly reduce the level of informal employment. These findings emphasize the importance of spatially-aware policy interventions in addressing the issue of informal employments. Strengthening education, digital infrastructure, and targeted poverty alleviation programs are important strategies.

***Keywords**: informal employment, poverty, education, internet access, spatial econometrics, Spatial Error Model*

INTRODUCTION

Informal activities have been recognized as a common phenomenon in many developing countries. Involvement in the informal economy is also associated with lower wages, more difficult and dangerous work, less security and protection, and many other disadvantages (Sibagariang et al., 2023). On the other hand, the informal economy remains a major source of livelihood, especially in developing countries, as it is able to provide immediate protection from poverty. In the absence of a social protection system for poverty, the informal sector acts as a quick solution to people's need for income, especially for people with low education (Mazlan et al., 2021).

This phenomenon can be explained through the dualism theory approach, which explains the structural division in the labor market between the formal and informal sectors. According

to Witjaksono in the research of Armelly et al. (2021), the dualism theory of the labor market also explains the segmentation based on the level of wages and benefits obtained. The formal sector offers high wages and a large return on investment in human capital, although the number of jobs available is very limited. In contrast, the informal sector offers low wages and is an alternative for less-educated workers and those who are forced to leave the formal sector due to labor rationalization. In line with this theory, Rusmana et al. (2025) revealed that the labor dualism theory defines informal sector workers as identical to characteristics such as low production standards, unstable income, and minimal job security. This condition has the potential to cause irregularities in the economic structure.

In Indonesia, the informal sector continues to play a dominant role in shaping the structure of employment and the broader economy. According to data from the Central Bureau of Statistics (BPS, 2024), approximately 59.11% of the country's workforce in 2023 was engaged in informal employment. This figure underscores a significant reliance on jobs that are often characterized by low productivity, limited legal protections, and the absence of formal employment benefits. Informal sector workers typically operate outside the scope of labor regulations, leaving them vulnerable to unstable incomes, poor working conditions, and exclusion from essential services such as health insurance and pension schemes. The high prevalence of informality is particularly evident in sectors such as agriculture, trade, transportation, and domestic services, where informal arrangements are often normalized. A study by Syafitri et al. (2024) reveals that individuals working in the informal sector frequently encounter structural barriers in accessing social protection systems and healthcare services. This lack of institutional support contributes to a cycle of economic insecurity and social vulnerability, particularly among low-income households. Furthermore, the informal nature of their employment often limits their ability to access financial services such as credit or savings programs, which are crucial for building economic resilience.

Reliance on the informal sector is often associated with the socio-economic conditions of a region. A study conducted by Sharma & Adhikari (2020) shows that poverty is the main factor that drives individuals to work in the informal sector, where limited economic resources leave people with no choice but to work in a more flexible but unprotected sector. This is supported by research conducted by Sibagariang et al. (2023) which states that the level of poverty has a significant effect in increasing the percentage of informal workers. In

addition, the low level of education completion also contributes to the high number of informal employments, as individuals with limited education have barriers in accessing formal jobs that require higher skills (Pulungan & Haryanto, 2024). Other research also shows that individuals with higher education tend to have a greater chance of working in the formal sector compared to the informal sector (Mkulisi & Sathiya Susuman, 2024).

In addition to poverty and education, the development of digital infrastructure also plays a role in the employment structure. Access to the internet plays a role in maintaining labor stability in the formal sector and reducing the risk of transition to the informal sector (Tasmilah et al., 2023). However, in reality, the use of the internet can increase the income of informal workers in both urban and rural areas, thus encouraging an increase in informal sector workers (Faisal & Rahadian, 2023). This shows that there are different perspectives on the effect of internet use on the level of informal employment.

Most of the previous studies still focus on the direct relationship between socio-economic factors and informal labor without taking into account the spatial aspect. In fact, the socio-economic condition of a province can have a spillover effect on other provinces around it (Fahmi, 2022; Ichsan et al., 2022). For example, the poverty rate in a region or province affects the poverty rate in its surrounding areas, which in turn has an impact on the distribution pattern of informal labor (Rahmadeni, 2020; Sari & Amalia, 2021). Regions with high poverty rates generally have limited access to formal employment, which encourages more of their residents to work in the informal sector or even migrate to other regions in the hope of obtaining more decent work (Mallett, 2018). Therefore, an approach that considers spatial linkages is crucial to understanding the dynamics of informal workers more comprehensively. To fill the gap of previous research, this study adopts the Spatial Error Model (SEM), which allows a more in-depth analysis of the spatial effects in the distribution of informal workers in Indonesia. The choice of this model is based on its ability to consider spatial autocorrelation in the residual variables, thus capturing patterns of interrelationships between regions that cannot be explained by ordinary regression (Nabila, 2025). Thus, this study aims to examine the socio-economic factors that influence the level of informal employment by considering area effects and spatial linkages between provinces in Indonesia.

This research is expected to generate more effective policy implications, especially in reducing dependence on the informal sector and encouraging a more stable transition to the

formal sector. If poverty is proven to be the main factor driving informal employment, then strengthening region-based poverty alleviation policies is necessary. On the other hand, if internet access plays a significant role in reducing the number of informal workers, developing digital infrastructure in disadvantaged areas could be an effective solution (Banga & Willem, 2023). In addition, if education has a large influence, then improving access and quality of education is a top priority in employment policy. By understanding the spatial linkages between regions, employment policies can be structured more strategically to create a more inclusive and sustainable work environment for all workers in Indonesia.

METHOD

This research uses a quantitative approach with spatial regression methods to analyze the socio-economic factors influencing the rate of informal workers across 34 provinces in Indonesia in 2023. Spatial regression is utilized to identify spatial dependence between regions, which cannot be explained by the classical linear regression model (Oktaviani, 2018). The study begins by testing for spatial autocorrelation using Moran's I, followed by the Lagrange Multiplier (LM) test to determine the most appropriate spatial model between the Spatial Lag Model (SLM) and the Spatial Error Model (SEM). These steps are based on research conducted by Novitasari & Khikmah (2019). Suppose the LM (error) test result is significant. In that case, SEM is deemed more appropriate, as this model captures the influence of unobserved factors in one region that may affect other regions (Japany & Firnanda, 2022). The SEM model is expressed in the following equation:

$$Y = X\beta + u, \quad u = \lambda W u + \varepsilon$$

Where Y is the dependent variable (informal sector employment rate), then X is the independent variable (percentage of poor population, high school completion rate, percentage of households that have ever accessed the internet). Whereas β is the regression parameter, W is the spatial weighting matrix, λ is the spatial dependency parameter, and ε is the independent error term. According to Anselin, as cited in the study by Faujia & Subarkah (2022), spatial modeling is carried out using GeoDa software, which allows for model estimation as well as visualization of spatial relationships between regions through maps and graphs.

This study analyzes 34 provinces in Indonesia in 2023 as the units of analysis without employing sampling techniques, as the research is conducted on an aggregate scale and based on secondary data. The data used were obtained from Badan Pusat Statistik (BPS) through publications such as Sakernas and Susenas.

Table 1. Variable Definition

Variables	Definition	Source
The Level of Informal Employment	The percentage (%) of informal workers out of the total labor force in each province in Indonesia	BPS (2024)
The percentage of the poor population	The percentage of the population living below the poverty line serves as an indicator of the level of economic welfare in a province	BPS (2023)
The senior high school completion rate	The percentage of individuals who have completed upper secondary education	OECD (2019)
The percentage of households that have accessed the internet	Calculating the total percentage of households that reported having accessed the internet within the past three months, in both urban and rural areas	BPS (2024a)

RESULTS AND DISCUSSION

Spatial Autocorrelation

The classification in this study was determined using the natural breaks model, which is considered effective in representing the natural distribution patterns of the data. This model divides the data into groups based on inherent distribution patterns. According to Rifaldi et al. (2024), this method is particularly effective when the data are unevenly distributed, as it allows for grouping without being biased toward any specific distribution shape. Figure 1 presents the natural breaks map of Indonesia for the year 2023, which includes the level of informal employment, the percentage of the poor population, senior high school completion rates, and the percentage of households that have accessed the internet.


Natural Breaks: INF
< 44.670 (2)
[44.670, 58.140) (10)
[58.140, 70.670) (17)
>= 70.670 (5)

(a)


Natural Breaks: PM
< 8.700 (17)
[8.700, 13.850) (9)
[13.850, 19.960) (5)
>= 19.960 (3)

(b)


Natural Breaks: SSMA
< 54.790 (3)
[54.790, 63.410) (6)
[63.410, 73.630) (17)
>= 73.630 (8)

(c)



(d)

Figure 1. Natural Breaks Map of Indonesia for 2023

(a) The level of informal employment; (b) The percentage of the poor population; (c) Senior High School Completion Rate; (d) Percentage of Households that Have Accessed the Internet

Based on the map in Figure 1, it was found that neighboring provinces have similar Natural Breaks classifications. This indicates that the observed values exhibit spatial dependence (Hikmah, 2017). However, to further confirm whether neighboring provinces have spatial relationships, a spatial autocorrelation test will be conducted using Moran's I. Moran's I test is used to analyze spatial relationships within an event across a region as a whole (Anuraga & Sulistiyawan, 2017). The value of Moran's I ranges from -1 to 1. A value below 0 indicates negative spatial autocorrelation, while a value above 0 indicates positive spatial autocorrelation. A value of 0 suggests no spatial autocorrelation (Mailanda et al., 2022). The results of spatial autocorrelation using Moran's I can be seen in Table 2.

Table 2. Results of Moran's I Test

No	Variable	Result of Moran's I
1	The level of informal employment	0.406
2	The percentage of the poor population	0.730
3	The senior high school completion rate	0.105
4	The percentage of households that have accessed the internet	0.242

The results of the Moran's I test presented in Table 2 show that the variable for the level of informal employment has a value of 0.406, indicating a moderate level of spatial autocorrelation. This suggests that the distribution of informal workers in a given area is influenced by neighboring regions. The percentage of the poor population has the highest Moran's I value at 0.730, indicating a strong spatial pattern. This means that areas with high poverty levels tend to be geographically close to other high-poverty areas.

In contrast, the senior high school completion rate has a Moran's I value of 0.105, reflecting a weak and more randomly distributed spatial pattern. Meanwhile, the percentage of households that have accessed the internet has a Moran's I value of 0.242, indicating some degree of spatial distribution, although not as pronounced as the patterns seen in poverty or informal employment. Overall, these results confirm that socioeconomic factors exhibit varying degrees of spatial interdependence, with poverty and informal employment being more strongly influenced by spatial effects.

Lagrange Multiplier (LM)

Before determining the appropriate spatial regression model, it is necessary to conduct an initial test of the possibility of spatial dependency effects in the data (Usali et al., 2021). One method commonly used for this purpose is the Lagrange Multiplier (LM) Test, as described by LeSage in Agustina et al. (2022). If the LM-lag value is significant, then the Spatial Autoregressive (SAR) model is selected. If the significant LM-error, then the Spatial Error Model (SEM) model is used. Therefore, in this study, the Lagrange Multiplier Test will be used as the basis for identifying the most suitable model, as shown in Table 3 below.

Table 3. Results of the Lagrange Multiplier Test

No	Uji	MI/DF	Value	Prob
1	Moran's I (error)	0.3881	3.078	0.002*
2	Lagrange Multiplier (lag)	1	0.065	0.798
3	Robust LM (lag)	1	0.091	0.762
4	Lagrange Multiplier (error)	1	6.594	0.010*
5	Robust LM (error)	1	6.621	0.010*
6	Lagrange Multiplier (SARMA)	2	6.686	0.035*

Based on the Lagrange Multiplier (LM) test results presented in Table 3, the model that incorporates spatial error effects is considered more appropriate for this study. The Moran's I (error) probability of 0.002 indicates the presence of spatial autocorrelation in the model, so that the OLS regression approach is not able to capture the spatial patterns that exist in the data. Furthermore, based on the LM (error) test results with a probability of 0.010 and Robust LM (error) of 0.010, indicating a significant spatial error effect. This indicates that the use of the Spatial Error Model (SEM) is more appropriate than the Spatial Lag Model (SLM). This is supported by the results of the LM (lag) test results, which show a probability of 0.798 and Robust LM (lag) of 0.762. Both results are insignificant, so the spatial lag effect does not dominate in this model. Therefore, the SEM model was chosen because it can capture the influence of unobserved factors that are spatially dispersed in explaining the level of informal employment in Indonesia.

Spatial Error Model (SEM)

The results obtained from the Lagrange Test then the SEM model is chosen because it is considered more appropriate. Where SEM is a Spatial Error Model used when the independent variables do not affect each other spatially, but the residuals show a spatial relationship between regions. This model takes into account spatial autocorrelation in the error term (Rahmawati & Bimanto, 2021). So we conducted the Spatial Error Model (SEM) Test with the results listed in Table 4.

Table 4. Results of Spatial Error Model Test (SEM)

Variable	Coeff	Std. Error	z-value	Prob
R-Squared	0.7625			
Prob. Breusch-Pagan Test	0.7111			
Prob. Likelihood Ratio Test	0.001*			
Constant	115.31	11.66	9.882	0.00*
The percentage of the poor population	0.508	0.268	1.892	0.05*
The senior high school completion rate	-0.289	0.101	-2.872	0.00*
The percentage of households that have accessed the internet	-0.482	0.116	-4.153	0.00*
Lambda	0.543	0.127	4.274	0.00*

The model evaluation results displayed in Table 4 show an R-squared value of 0.7625, which means that about 76.25% of the determinants of the level of informal employment can be explained in the model. In addition, the probability value of the Breusch-Pagan test is 0.711, which means that there is no heteroscedasticity problem in the model, so the assumption of homoscedasticity can be accepted. Meanwhile, the Likelihood Ratio test results show a probability of 0.001, which means that the Spatial Error Model (SEM) is more appropriate than the ordinary OLS regression model due to the presence of spatial autocorrelation in the error model. This result indicates that the distribution of informal employments is not only influenced by factors within an area but also by conditions in the surrounding area.

Based on the Spatial Error Model (SEM) estimation results in Table 4, it is known that socio-economic factors have a significant influence on the level of informal employment in Indonesia. The probability value of 0.00 at the 5% significance level indicates a significant influence, which means that even though the independent variable is zero, other factors still affect the level of informal employment. This result confirms that the phenomenon of informal workers in Indonesia is a complex issue that is not only influenced by the variables in the model but also by other factors that are not observed in this study.

The percentage of poor population variable has a coefficient of 0.508 and a probability of 0.05 with a significance level of 5%, which indicates that the result is significant. This means that every 1% increase in the poverty rate in a region will increase the informal employment rate by 0.508%. This result is consistent with the study by Sharma & Adhikari (2020) and Lahengko & Suryono (2023), which states that the informal sector often functions as a “buffer zone” for workers who cannot be absorbed in the formal sector due to limited employment opportunities and high poverty rates. In Indonesia, areas with higher poverty rates tend to have limited access to formal jobs, thus encouraging people to find jobs in the informal sector, which is more accessible, even with more unstable working conditions.

On the other hand, the high school completion rate has a negative coefficient of -0.289 and a probability of 0.00 with a significance level of 5%, which means the result is significant. This suggests that an increase in secondary education completion can reduce the level of informal employment. This result supports the Human Capital theory, which states that an increase in education level contributes to an increase in skills and labor productivity, thus

opening up greater opportunities to obtain employment in the more stable formal sector. The study by Psacharopoulos & Patrinos (2018) shows that each additional year of education significantly increases individual income, reflecting the importance of investment in human capital. Higher education allows individuals to access jobs with better contracts, higher salaries, and better labor protection.

The variable percentage of households that have accessed the internet significantly affects the level of informal sector workers. With a coefficient of -0.482, it indicates that every one percent increase in the level of households that have accessed the internet will reduce the level of informal employment by 0.48%. This finding is in line with a study conducted by Tasmilah et al. (2023), where the use of the internet by formal employments helps maintain the existence of formal jobs and reduces the possibility of switching to the informal sector. The internet serves to improve digital skills, create new job opportunities in the formal sector, and expand access to information about formal job vacancies.

In addition, the lambda coefficient of 0.543 significantly affects the level of informal employment, indicating a significant impact of spatial error in this model. This suggests that there are factors not included in the model that affect the level of informal employment in a region while also affecting conditions in neighboring provinces. This finding reinforces the concept of spatial spillover effects in regional economics, where inequality between regions can increase or decrease the number of informal workers in a region, depending on the economic and social conditions of the surrounding areas (Chocholatá & Furková, 2023; Garcia et al., 2024).

CONCLUSION

The results of this study underscore that enhancing educational attainment and expanding access to digital technology are key strategies in reducing informal employment in Indonesia. Higher levels of secondary education completion and greater internet access are associated with lower rates of informal work, reflecting the importance of investing in human capital and digital inclusion. At the same time, poverty remains a significant driver of informal employment, highlighting the urgent need for sustained poverty reduction efforts. Furthermore, the presence of spatial spillover effects indicates that regional disparities influence labor market dynamics across provinces. Therefore, policy interventions must be

spatially sensitive and regionally coordinated to effectively address the structural roots of informality and promote more inclusive and formal labor markets nationwide.

As the recommendation for future research, it is recommended to explore other socio-economic variables that may influence the level of informal employment, such as the level of urbanization, infrastructure quality, income inequality, or gender involvement in the world of work. In addition, the use of more complex spatial methods such as Spatial Durbin Model (SDM) or Geographically Weighted Regression (GWR) can also be considered to capture the influence of location in more detail and dynamically. This approach will enrich the understanding of informal sector dynamics and provide more comprehensive inputs for area-based policy formulation.

REFERENCES

387-JEBE-Similarity- --

ORIGINALITY REPORT

10%	4%	8%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Colin C. Williams, Jan Windebank. "Informal Employment in the Advanced Economies - Implications for Work and Welfare", Routledge, 2002 Publication	2%
2	www.mdpi.com Internet Source	1%
3	Shanaka Herath. "Discovering the Urban Structure Using a Spatial Hedonic House Price Model", ICCREM 2013, 2013 Publication	1%
4	Submitted to University of Florida Student Paper	1%
5	Xiao-Ling Song, Ya-Ge Jing, Kade'erya Akeba'erjiang. "Spatial econometric analysis of digital financial inclusion in China", International Journal of Development Issues, 2020 Publication	1%
6	assets-eu.researchsquare.com Internet Source	1%
7	Robert Fornango. "When Space Matters: Spatial Dependence, Diagnostics, and Regression Models", Journal of Criminal Justice Education, 06/2010 Publication	1%

8	link.springer.com Internet Source	<1%
9	docu.tips Internet Source	<1%
10	repository.uki.ac.id Internet Source	<1%
11	Imelda K. Moise, Ezekiel Kalipeni. "Applications of geospatial analysis to surveillance data: a spatial examination of HIV/AIDS prevalence in Zambia", GeoJournal, 2010 Publication	<1%
12	KWANG-KOO KIM. "Natural Amenities and Rural Development: Understanding Spatial and Distributional Attributes", Growth and Change, 6/2005 Publication	<1%
13	ebin.pub Internet Source	<1%
14	www.frontiersin.org Internet Source	<1%
15	Gebremeskel H. Gebremariam, Tesfa G. Gebremedhin, Peter V. Schaeffer. "County-level determinants of local public services in Appalachia: a multivariate spatial autoregressive model approach", The Annals of Regional Science, 2011 Publication	<1%
16	Akio Kusakabe. "SPATIAL ANALYSIS OF THE JAPANESE GAS INDUSTRY", Review of Urban and Regional Development Studies, 07/2010 Publication	<1%

17	Zhang, T.. "A decomposition of Moran's I for clustering detection", Computational Statistics and Data Analysis, 20070815 Publication	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		