university of groningen

faculty of spatial sciences

Master's Thesis

SimJogja: A Spatial Microsimulation Approach to COVID-19 Vulnerability and Impact Assessment of the Local Labour Market

Immanuel Satya Pekerti
S4618602

Under supervision of
Prof. Dimitris Ballas

A thesis submitted as partial fulfilment of the requirements for the degree of Master of Science in Economic Geography: Regional Competitiveness and Trade

Faculty of Spatial Sciences
University of Groningen

2023

Foreword

This thesis is another milestone in my progress of learning. The successful completion of my thesis marks the end of my master's programme at the University of Groningen but marks another start in my lifelong journey of learning. Exploring the *terra incognita* of our knowledge has always been challenging. For me, it was arduous and confusing at most times. I have no better reference to the struggle I went through than this almost philosophical quote in Indonesian:

> "Memang mbak nggak bingung? Semua manusia di muka bumi ini bingung, mbak. Nanti nggak bingung kalau udah di surga." -Aldi Taher
> (Are you not confused? Every one on this earth are confused. Only in heaven we'll be free of the confusion)

Despite the struggle, the journey was worth it. I am very thankful for the support bestowed by my supervisor, Prof. Dimitris Ballas, my mother, my dearest one, and also my colleagues and friends.
I also hope that this thesis serves a bigger purpose for society than merely sitting idle in the archive.

Bogor, August 30, 2023

Immanuel Satya Pekerti

SimJogja: A Spatial Microsimulation Approach to COVID-19 Vulnerability and Impact Assessment of the Local Labour Market

SUBMITTED VERSION

Immanuel Satya Pekerti*

Abstract

I build a static spatial microsimulation model, called SimJogja, to evaluate the labour market impact of COVID-19 in the Indonesian province of D.I. Yogyakarta. I argue that current literature overlooks the importance of geography in the investigation of COVID-19's economic impact. This paper adds to the literature by demonstrating the spatial dependence and heterogeneity through the estimation of population microdata on the village level, downscaling the currently available data and making available one with higher resolution. This paper is also expected to have policy relevance by demonstrating its capability to generate reliable and finely disaggregated simulated data to support geographic targeting, and what-if analysis to provide an ex-ante or ex-post local policy impact assessment. This is the first spatial microsimulation model for Indonesia and the first known use of spatial microsimulation to assess the socio-economic and labour market impact of COVID-19.

Keyword : spatial microsimulation; COVID-19; local; labour market; Yogyakarta

*Faculty of Spatial Science, University of Groningen, i.s.pekerti@student.rug.nl

Contents

1 Introduction

Pandemics have recurred across history. Even so, the COVID-19 pandemic came to us as a surprise, a new kind of pandemic that virtually no one understands how to handle when it appeared. This pandemic has impacted various facets of our society, including the economy. Academic communities have strived to understand the nature of its impact, and governments have tried to protect their economy along with the livelihood of their constituents. The apparent struggle with the COVID-19 pandemic shows that a lot needs to be done. My research highlights the need for high-quality, high-resolution economic data to fulfil various academic and policy purposes to face such an adverse shock through a geographical approach, and proposes spatial microsimulation to answer the need.

The urgency of this paper stems from the gap in the current academic literature on COVID-19 economic impact, especially on the labour market. I found that the current literature focuses more on 'who' (Adams-Prassl et al., 2020; Aum et al., 2021) with too few regards to 'where'. Even if some discuss the 'where', it often stops at urban-rural dichotomy (for example Lee et al., 2021). I argue that geography plays more role than can be captured with such a framework. Older spatial analysis literature, such as Rey and Montouri (1999), already recognises more complex economic interaction patterns between places, such as substantive and nuisance spatial dependence. Alas, these patterns won't be so obvious given the current level of disaggregation of geographic data.

Recently available economic data in Indonesia, especially the ones provided by *Badan Pusat Statistik* (Indonesian Central Statistical Bureau/BPS), are mostly geocoded at the district level. According to the BPS, the average area of districts in D.I. Yogyakarta is 626.63 km squared. This degree of generality makes it hard to observe the geographic patterns that emerge from various economic variables.

Another motivation is policy-related. The pandemic threatened various vulnerable demographic groups and hit some groups disproportionately harder than others (Adams-Prassl et al., 2020; Lee et al., 2021). The government's role in mitigating the impact and assisting those who are in need is essential to protect their livelihood and the economy as a whole from a severe downturn. Geographic targeting is one of the assistance targeting measures which has a potential for improvement. The accuracy of the targeting measure could be improved with higher resolution mapping (Baker and Grosh, 1994; Elbers et al., 2007).

Both dimensions of motivation could benefit from the generation and estimation of high-quality, high-resolution geographic data. Hence, I'm proposing spatial microsimulation as the tool to fulfil this objective. This paper is an attempt to demonstrate how the method could achieve the aforementioned objectives.

This paper describes the *SimJogja*, a static spatial microsimulation model of the local labour market in the Indonesian province of D.I. Yogyakarta. Furthermore, I demonstrate how it helps evaluate the impact of COVID-19 through geographic information system (GIS) visualisation and simple what-if analysis. SimJogja is not the first spatial microsimulation model to to analyse the local labour market. SimLeeds (Ballas et al., 2006a; Ballas and Clarke, 2000) and SVERIGE (Rephann et al., 2005) have also been used to examine the local labour market impact through simulation of a hypothetical scenario or estimation of a small area population, but the extent of my knowledge, SimJogja is the first one to apply this method on the context of COVID-19.

I expect this paper to fulfil three outcomes with both academic and policy relevance. First, to demonstrate how to generate high-resolution geographic data that could help boost the accuracy of geographic targeting. Second, to illustrate the usefulness of what-if analysis for the policymaking process in the COVID-19 context. Third, highlight the geographic pattern of the labour market outcomes and set the stage for future research on the role of geography in local labour market outcomes in the context of negative economic shock. Such knowledge enables us to better understand the important nexus between geography, health, and economic outcomes.

I use the case of an Indonesian province, *Daerah Istimewa Yogyakarta* (Special Region of Yogyakarta/D.I. Yogyakarta). This decision is made considering the region's reliance on the tourism sector, one of the hardest hit by the pandemic. Sun et al. (2021) estimated 242,000 jobs were vulnerable due to the pandemic, second only to Bali with 820,000 jobs vulnerable. Moreover, the province made publicly available various high-quality and detailed demographic information at the village level, which is essential to the reliability of the spatial microsimulation model.

Figure 1: Reference Map

2 COVID-19 and Labour Market: Making the Case for Geography

In this section, I argue that the current body of literature doesn't talk enough about the role of geography in the investigation of COVID-19 economic impact. I discuss two important points. First is how current literature focus on 'who' and understates the importance of 'where' as a complement. Additionally, I highlight the lack of geographical analysis of various labour market outcomes and argue for its urgency. Second is how urban-rural dichotomy, incorporated in various research frameworks of current literature, is not enough to capture the full extent of geography's role in this issue. Through this discussion, I point out the academic gaps that this paper intends to fill.

Space is an important dimension in both economic and health analysis, and yet, it did not receive the amount of attention it deserves in the investigation of COVID-19 impact. A larger portion of the literature is concerned with identifying the most vulnerable groups and the disproportionate impact they faced (Adams-Prassl et al., 2020; Alon et al., 2020; Fairlie et al., 2020; Lee et al., 2021; Singh et al., 2022), how use of internet in jobs promote employment resilience (Adams-Prassl et al., 2022, 2020; Angelucci et al., 2020; Dingel and Neiman, 2020), and the analysis of labour market transition for each of employment categories in response to the shock (Cowan, 2020). These topics are undoubtedly important, but leaving out geography also leaves out the essential part of the picture. Many aspects of the labour market are geographic, such as job accessibility (Andersson et al., 2018), spatial sorting (Mion and Naticchioni, 2009), and local impact of job loss (Ballas et al., 2006a).

The problem at hand is to put together pieces of evidence to unveil the nexus between space, economic, and health outcomes within the COVID-19 context. The nexus is evident and well-documented. Literature has documented evidence of spatial inequality of health (Ballas et al., 2006b; Broomhead et al., 2023) and the relationship between economic inequality and health outcomes (Zhao et al., 2021). Geographical disparity of labour market conditions is also evident through a study by López-Bazo et al. (2005) who studied and proved the geographical distribution of unemployment in Spain.

Within the COVID-19 context, there are also several studies investigating different sides of the relationship. They are often interested in how a combination of the two factors affect the other: how economic activity and pandemic shock shape the geography (Delventhal et al., 2022; González-Leonardo et al., 2023; Ramani and Bloom, 2021), how the spatial concentration of economic condition or mobility determines the health outcomes (Desmet and Wacziarg, 2022), or how economic resilience comes about through geographic concentration in the face of the pandemic (Dai et al., 2021).

Given the literature, I identify an important gap. That is how health-related shocks affect economic outcomes, specifically the labour market, across places. I showed earlier that literature has managed to identify groups most susceptible to the pandemic and the resulting labour market outcomes, but identification of their geographical distribution is severely lacking. The insight into the distribution can shed light on the probable cause of the spatially heterogeneous impact. Inequality between places is sufficiently evident, but our understanding of its cause is rather insufficient.

Geography and distance, left unbridged, can isolate and alienate some unfortunate groups from the available economic opportunities located elsewhere. This is well manifested in the First Law of Geography (Tobler, 1970) and the transport cost of the New Economic Geography model (Krugman, 1991). The consequence is there will be places that don't matter as much as others, as Rodríguez-Pose (2018) has argued. Much is left to be desired for a socio-spatially just society. Hence the urgency to understand better the problem in the geographical distribution of labour market outcomes.

I now raise the second issue of the shortcomings of the urban-rural framework. The urban-rural framework adopted in various analyses in the existing literature is too simplistic. Its use is sufficient to look for the magnitude of difference between areas assigned as rural and urban, but it doesn't allow us to look deeper into the geographical process at play. I point out two major issues with this framework. First, it overlooks the other, more important spatial effects. Rey and Montouri (1999) suggest that there are two types of spatial effects: spatial dependence and spatial heterogeneity. The spatial dependence is further differentiated between substantive and nuisance dependence. Substantive dependence refers to the difference between regions due to their characteristics, and the urban-rural dichotomy is an example. Nuisance dependence refers to the fact that economic and administrative boundaries might not correspond to each other. For example, a dense economic activity might happen just around the administrative border. In the case of D.I. Yogyakarta, this happens around the Sleman-Yogyakarta border. Finally, spatial heterogeneity refers to the varying pattern of relationships between points across space. The current framework can't capture nuisance dependence and spatial heterogeneity. A new approach is needed to incorporate these effects.

Second, the definition of urban areas. COVID-19 hits harder where people are closer together, and urban areas, given their density, are put in the limelight across COVID-19 studies. Despite that, the identification of urban areas is somewhat problematic. The problem here follows from the nuisance dependence that we have discussed earlier and is a concern in previous studies concerning the urban living condition (Henderson and Turner, 2020). As an example, we have two given facts in the case of the Sleman-Yogyakarta border: (i) Sleman is predefined as a rural area by the Badan Pusat Statistik (Indonesian Central Statistical Bureau/BPS), and (ii) parts of Sleman bordering Yogyakarta are urbanised. The implication will be that the economic impact resulting from the pandemic will be different for each part of Sleman, and this is lost in the single parameter estimated by econometric specification in the form of an urban dummy. I demonstrate later that spatial microsimulation can downscale the data and alleviate this problem.

Additionally, I also demonstrate how spatial microsimulation can evaluate the spatial heterogeneity of a scenario impact through what-if analysis.

3 The Vulnerable Groups and Indicators of Labour Market Transitions: Global and Indonesian Context

In this section, I discuss the relevant literature to correctly identify the groups and indicators that are related to the pandemic to provide direction for my analysis. Studies have found various demographic groups that are hardest hit during the pandemic: women, youth, the less educated, racial minority, informal workers, workers with temporary contracts, low-skilled and low-paying jobs, and jobs with tasks that can't be done remotely (Adams-Prassl et al., 2020; Aum et al., 2021; International Labour Organization, 2020; Lee et al., 2020; Lee et al., 2021). The impact of COVID-19 also hit some sectors hard, such as tourism and leisure activity (Manning, 2021; Wren-Lewis, 2020). I decided to focus on four groups: informal workers, workers in the tourism sector, along with women and youth workers. These groups are chosen based on different notions of vulnerability. Informal workers are chosen due to their inability to withstand shocks (International Labour Organization, 2020). Workers in the tourism sector are worth highlighting since they are exposed to the possible employment effect of output decline as was suggested by Manning (2021). Women and young workers are vulnerable due to their over-representation in vulnerable industries or jobs with unfavourable characteristics (e.g. low-skill, low-paying jobs, temporary contract arrangement, etc.) (Adams-Prassl et al., 2020; Lee et al., 2020).

Regarding the labour market outcomes, Manning (2021) suggests that there is no single reliable indicator. He points out that, while unemployment is proper to use in urbanised and developed economies, it might not be true for the case of developing countries. In concordance, Lee et al. (2020) suggests that unemployment is not a proper indicator of work disruption, as there are numerous cases of workers being stood down or temporarily not working, reduction of working hours, or being forced to exit the labour force altogether (see for example Coibion et al., 2020; Lemieux et al., 2020; Leyva and Urrutia, 2023). The share of informal workers is another interesting indicator to follow. I discuss some of these variables further.

The share of informal employment is essential given its high share in Indonesia. The share for Indonesia and D.I. Yogyakarta, was 55.88 per cent and 51.64 per cent in 2019, respectively. Any shock to informal jobs will have a widespread and significant effect towards the economy.

The debate on COVID-19's impact towards informality extends the importance of this measure. Lee et al. (2020) suggest that informal work was usually the last-resort option of the thrown-out formal workers, but is now no longer an option due to the mobility restriction. This is well-aligned with the findings of Leyva and Urrutia (2023) who find that decline in employment also sees a decline in informality rate due to informal productivity shock. On the contrary for the Indonesian case, Basri and Fitrania (2022) and Warr and Yusuf (2021) find a rise in the number of informal workers in the form of unpaid family workers and casual or unskilled farmers for the Indonesian case.

Another important thing is the association between working hours and informal jobs. Manning (2021) pointed out that reduction in working hours is a phenomenon more prevalent among informal workers. This might be one of the labour-saving strategies employed by firms to survive. Hence the need to examine the working hours reduction along with the informality rate.

Changes in the unemployment rate and labour force participation rate also warrant a close examination. Early examination by Coibion et al. (2020) shows that the decline in employment is not mirrored by an increase in the unemployment rate, but by a decrease in the labour force participation rate. This is confirmed by Aum et al. (2021). They guess that this is due to the unemployed waiting out the pandemic to get a better chance on job search, or expecting to go back to their previous employment. Coibion et al. (2020) points to early retirement given the raising figure of retired persons and workers declaring out of the labour force to retire.

4 Data and Methodology

4.1 Data Description

The model will use *Survei Angkatan Kerja Nasional* (Indonesian National Labour Force Survey/SAKERNAS) and the village statistics provided by the Provincial Government of D.I. Yogyakarta[1]. The scope of data will include villages and workers who reside in the province of D.I. Yogyakarta during each of the survey periods. There are 438 villages within my dataset, along with 9,249 observed individuals in 2020 and 9,414 individuals in 2019. No migration is assumed to avoid unnecessary overcomplication for now.

4.2 Brief Introduction to Spatial Microsimulation

Spatial microsimulation is one of the small area estimation methods, which concerns the production of reliable estimates of geographical units' characteristics with small or no samples (Pfeffermann, 2013). Spatial microsimulation is a development from its original, aspatial version, developed by Orcutt (1957) to model demographic changes, predicting the aggregate changes based on individual-level decision making. Microsimulation is concerned with the estimation of population microdata with a particular purpose of local-level policy impact analysis (Ballas et al., 2005). The idea to extend this by incorporating geography was conceived by Wilson and Pownall (1976) and put into action by Clarke et al. (1984). The method has seen various development in concept and application along with the development of a more advanced and powerful computing system. The main advantage of spatial microsimulation compared to other traditional small area estimation methods is its ability to construct whole synthetic microdata, or made-up data, for each geographical unit (Tanton, 2013). This enables us to go beyond a point estimate of each variable and get deeper insights from the cross-tabulation of variables of interest. Spatial microsimulation is also a way to add geography, even on a more micro scale, to the existing individual-level survey data (Ballas et al., 2005; Lovelace and Dumont, 2016). Further readings on this can be

[1] Data were retrieved from https://kependudukan.jogjaprov.go.id/

found in Ballas et al. (2005) for conceptual matters, and Smith et al. (2021) or Birkin and Wu (2012) for a review of its application in health and public policy, respectively.

There are a variety of ways for spatial microsimulation to work. Spatial microsimulation can generate synthetic microdata with aggregate characteristics that match the reference dataset (e.g. demographic statistics), which we call synthetic reconstruction. It could also assign a new weight to each observation for each of the geographical units, which we call reweighting. The principle of reweighting works by adjusting how much an individual observation with a unique set of characteristics is represented in a geographical unit. There are also two different ways of assigning new weights according to their process: probabilistic and deterministic. Tanton (2013) provides a complete review of the available models.

This study uses a deterministic reweighting approach, namely the iterative proportional fitting (IPF). IPF involves two types of variables: constraint and target. Constraint variables appear both in small area tables and the survey sample. They are usually more common variables that represent characteristics (e.g. age, educational attainment, gender, etc.). Target variables are usually only available on the survey sample dataset. They are usually more specific variables to be examined (e.g. unemployment status, earnings, etc.). Adjustment is made by calculating how likely a unique observation in the survey sample appear in a synthetic dataset based on the constraint variables value in the small area tables. This is done iteratively for each constraint variable, hence the name iterative proportional fitting. The accuracy of the synthetic dataset relies on the correlation between the constraint and target variables. In other words, it relies on how well a set of constraint characteristics could represent a set of target characteristics. I suggest readers consult Lomax and Norman (2016), which is also intended for nonexpert users, for a more extensive explanation of the way IPF works, especially on the iterative nature of the method.

4.3 Identification Strategy

This paper has three general aims: (i) to estimate the geographical distribution of the vulnerable groups and (ii) changes in labour market outcomes, and also (iii) simulate the geographical distribution of hypothetical scenarios' impact. The model-building process uses R software (R Core Team, 2023) and the `ipfp` package (Blocker, 2022) as described by Lovelace and Dumont (2016). I will discuss the identification strategy for each of the aims below.

Estimating the geographical distribution of the vulnerable groups is straightforward. I build a static spatial microsimulation model which maps the representation of the most vulnerable demographic groups across villages in D.I. Yogyakarta. The microdata sample used in this model is based on the August 2019 wave of SAKERNAS. This model aims to map the most vulnerable villages before the onset of COVID-19 in February 2020.

Identifying the geographical distribution of changes in labour market outcomes adds a bit more to the first model. I build another static model, representing the post-COVID outbreak. I have now two static models, with the first one based on the August 2019 wave of SAKERNAS, while the second one uses the August 2020 wave. Since there is no village proliferation or changes in the number of villages, finding the difference in variables value between the two periods is rather straightforward.

Simulation of hypothetical scenarios' impact is done through the what-if analysis. This simple analysis is performed by manipulating a value of a target variable for a part of population with certain characteristics, creating a post-treatment model. The difference between pre- and post-treatment is then mapped in GIS to observe the geographical distribution of the impact.

The constraint variables in this model are age, gender, age groups, occupational groups, and educational attainment. Each of the variables is divided into several categories. A complete list of these categories is available in the Appendix. The target variables are the variables that will be used to assess vulnerable demographic groups, such as working women, working youth, informal workers, and workers in the accommodation, food, and beverage industry.

I found a limitation in using SAKERNAS data in this model to identify the changes in labour market outcomes. World Bank (2020) found that around 70 per cent who lost their job early in the pandemic found their way back into employment by early August. The estimation I get by using the microdata might not show the full extent of the COVID-19 labour market impact.

4.4 Model Validation

The model will be useful only if it could reliably simulate the actual situation. Hence the need for validation. Validation is done by comparing actual and simulated values. Spatial microsimulation by reweighting is usually validated using internal and external validation.

Internal validation refers to the comparison of constraint variables' values between the input and simulated microdata, while external validation refers to the comparison of target variables' values between the population and simulated microdata. The result of the internal validation is shown in Figure 2. The internal validation compares the frequency of each category of the constraint variables between the simulated and actual population based on the small area tables used. The result shows a high correlation between actual and simulated values because the IPF conceptually works by optimising the fit of constraint variables between the two data (Panori et al., 2017). Despite that, my model found a considerable difference from two categories of education variable for each year. The largest difference between SAKERNAS and simulated data is found among those (i) who attain high school education and those (ii) who have not finished elementary school.

Edwards and Tanton (2012) suggest that external validation should be done against a different dataset. Restricted by data availability, I compare the mean of target variables aggregated value between simulated data with SAKERNAS survey data on the province level. Since the simulated weight and weight assigned by SAKERNAS went through two completely different processes, the SAKERNAS weight can be used for validation since its formulation is external to the estimation process (Edwards and Tanton, 2012).

The result of the external validation is provided in Table 1. I follow Panori et al. (2017) for the external validation procedure by looking at the difference in the average value for target

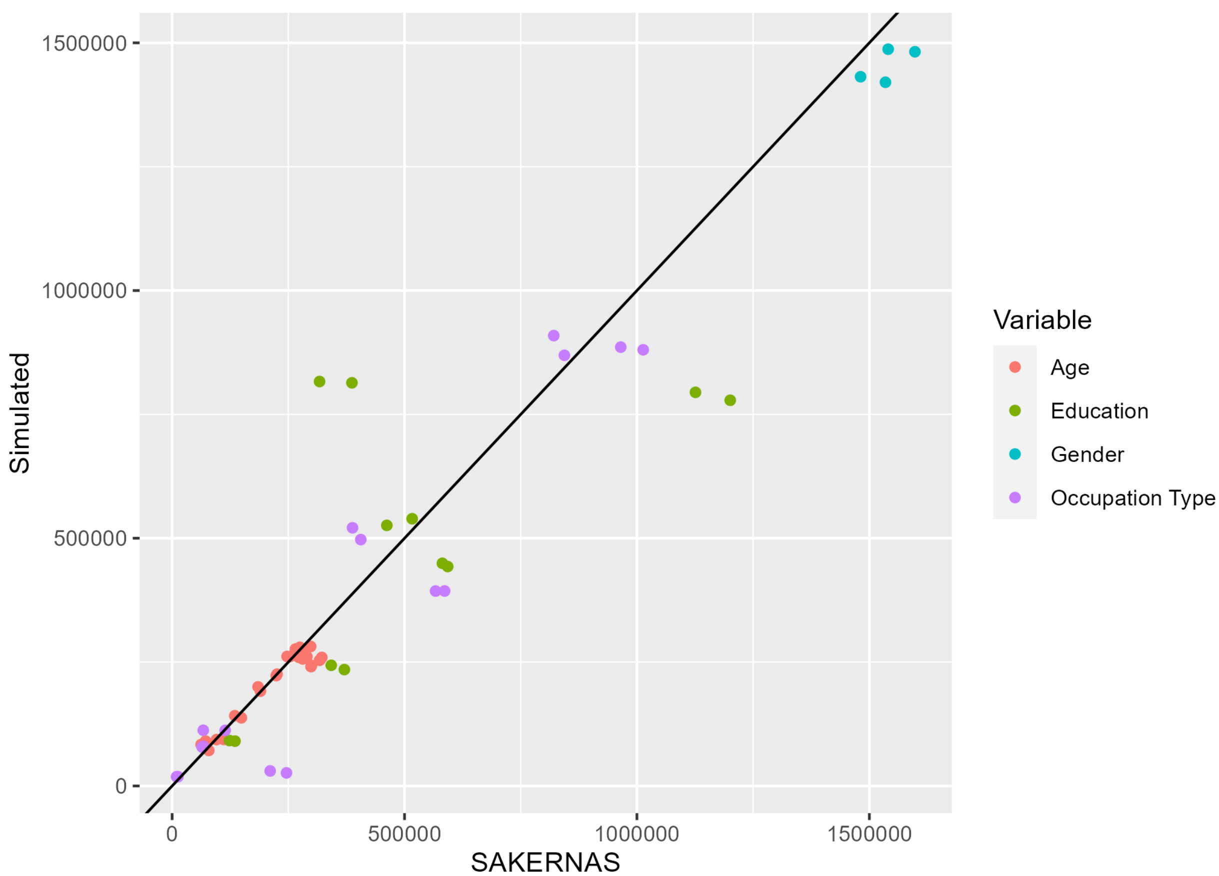

Figure 2: Scatterplot comparing the frequency of each category within the constraint variable, 2019-2020: Simulated vs. SAKERNAS

variables after aggregation. The table only displays variables which we will discuss. The validation of more target variables can be seen in the appendix.

To begin the discussion on the validation result, it is worth noting that there is no known well-defined rule of what constitutes an acceptable result. SimAthens' claim that their model is 'extremely satisfactory' is completely arbitrary. Therefore, I would not claim that my model is 'satisfactory' and provide my result for comparative purposes instead.

The largest difference is seen for the share of youth in employment, which is underestimated by 3.66 percentage points in the 2019 model. Another considerable difference is seen for the labour force participation rate, which is underestimated by 3.16 per cent in the 2019 model. The concern lies not in the mere difference, but in the magnitude of the time trend. The labour force participation rate dropped 5.18 percentage points between 2019 and 2020 using weight assigned by SAKERNAS, while it only dropped 0.81 percentage points using simulated value. This is particularly concerning since I'll be using the difference between both models. One might wonder if the difference in the share of youth employment would be even more alarming since it sees a change of direction between years. It is not a problem since the 2020 value of the variable won't be used in my analysis.

Table 1: External Validation of Target Variables

Variable	2019			2020		
	Simulated	SAKERNAS	Difference	Simulated	SAKERNAS	Difference
Working Hours	36.421	37.153	-0.73	34.796	35.340	-0.54
Unemployment Rate	5.54	3.04	2.50	5.59	5.38	0.21
Labour Force Participation Rate	69.67	72.83	-3.16	68.86	67.65	1.21
Share of Informal Workers	50.50	50.23	0.27	51.17	53.14	-1.97
Share of Women in Employment	41.75	43.61	-1.86	43.18	44.20	-1.01
Share of Youth in Employment	18.41	22.07	-3.66	19.47	21.15	-1.68
Sectoral Share: Accomodation & FnB	7.78	9.39	-1.60	6.47	8.24	-1.77

Note: Every variables are in percentage, except for working hours which is in average hours

The reason for the deviation might be the small number of samples (Edwards and Tanton, 2012). Spatial microsimulation relies on the inference of target variables based on the available combination of constraint variables' characteristics. A small selection of samples means that there are limited combinations of target variable characteristics that can be represented by a certain combination of constraint variable characteristics. In my case, I have 1,235 cases of working youth out of 9,414 samples used in the 2019 model. These samples are then used to infer the labour characteristics of 765,969 working youth out of the 2,905,051 working-age population.

5 Simulation Results

5.1 Estimation of Vulnerable Population

I discuss here the simulation result, and I start with the identification of the geographic distribution of vulnerable demographic groups. I select a few vulnerable categories to display (see Figure 3). The selected variables are the value of several categories as a share of employment in D.I. Yogyakarta for the year 2019. These categories are (i) share of informal workers, (ii) sectoral share of employment for accommodation, food, and beverage industry, (iii) share of women in employment, and (iv) share of youth (15-29 years old) in employment.

The share of informal workers is mostly concentrated outside of the urban core of Kota Yogyakarta. The concentration of high informality rate is found on the southern coast and northern part of Gunung Kidul and the Menoreh Hills on the western edge of the province, while the low informality rate is concentrated in the urbanised border area between the Kota Yogyakarta and Sleman. Interestingly, the share of women in employment somewhat mimics the distribution of informal workers. A high share of women's employment is concentrated on the southern coast of Gunung Kidul and Menoreh Hills. This might reflect the empirical finding that women are vulnerable to the pandemic due to its over-representation in informal or low-paying jobs (Aum et al., 2021).

Next is the share of employment in tourism-related industry. I use the accommodation and food, and beverage industry to represent the tourism industry in my analysis. The concentration of high employment rates in accommodation, food, and beverage is found around the urban areas and its immediate vicinity. Another area with a high share of employment is located around Wonosari. A lower level of employment share is present also on other known tourist spots, such as the base of Mount Merapi, Parangtritis Beach area, and Gunung Kidul beaches.

The share of youth employment is mostly concentrated around the urban area. Some sporadic concentration is also found on the northern edge of Gunung Kidul and a spot somewhere along its coast. The pattern does not correspond with any pattern of other variables. This might be associated with the correspondence between the type of jobs available in particular places and the type of jobs Indonesian youth would be associated with.

The general impression that one can conclude from these results is that the urban core and remote part of the province are both vulnerable, but in different ways. We see how tourism-related jobs and youth employment were concentrated mostly around the urban areas, while informal and women workers were mostly situated in more remote areas.

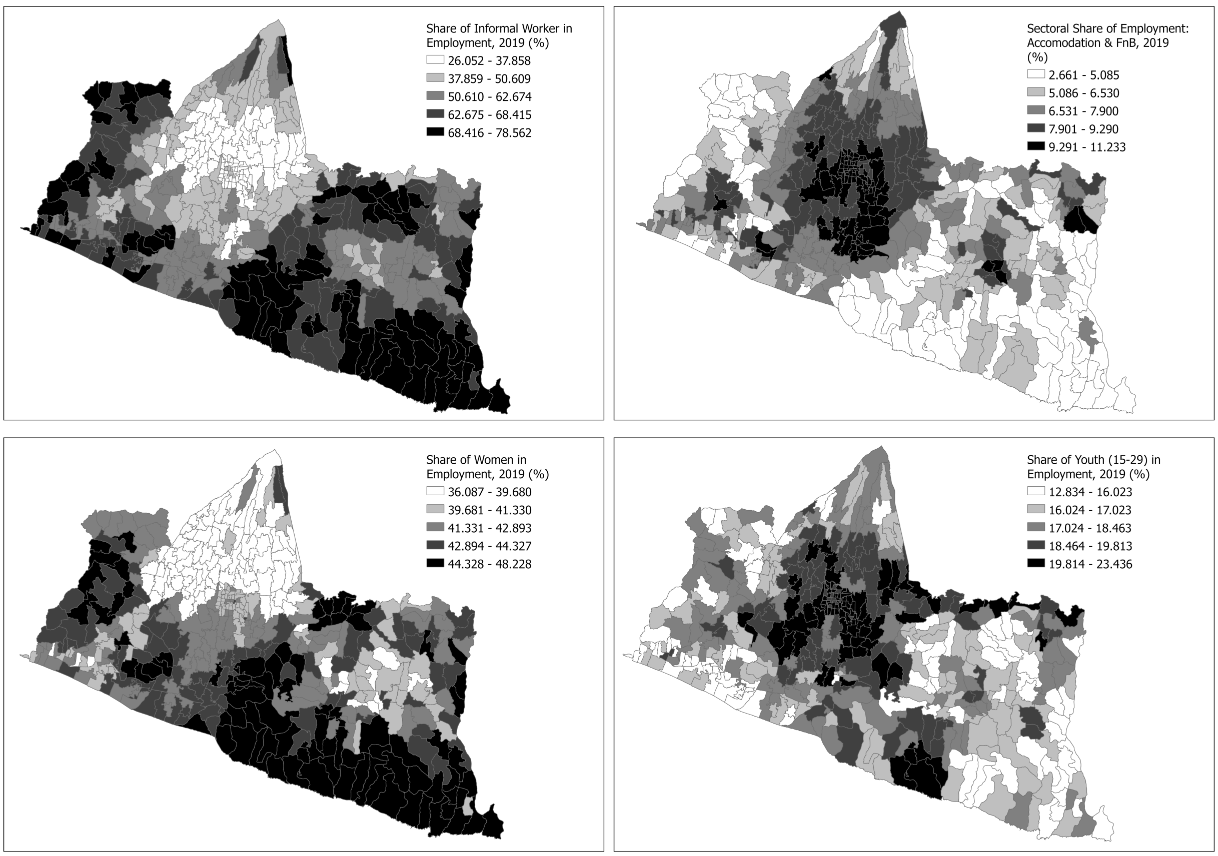

Figure 3: Geographical Distribution of Vulnerable Groups in 2019

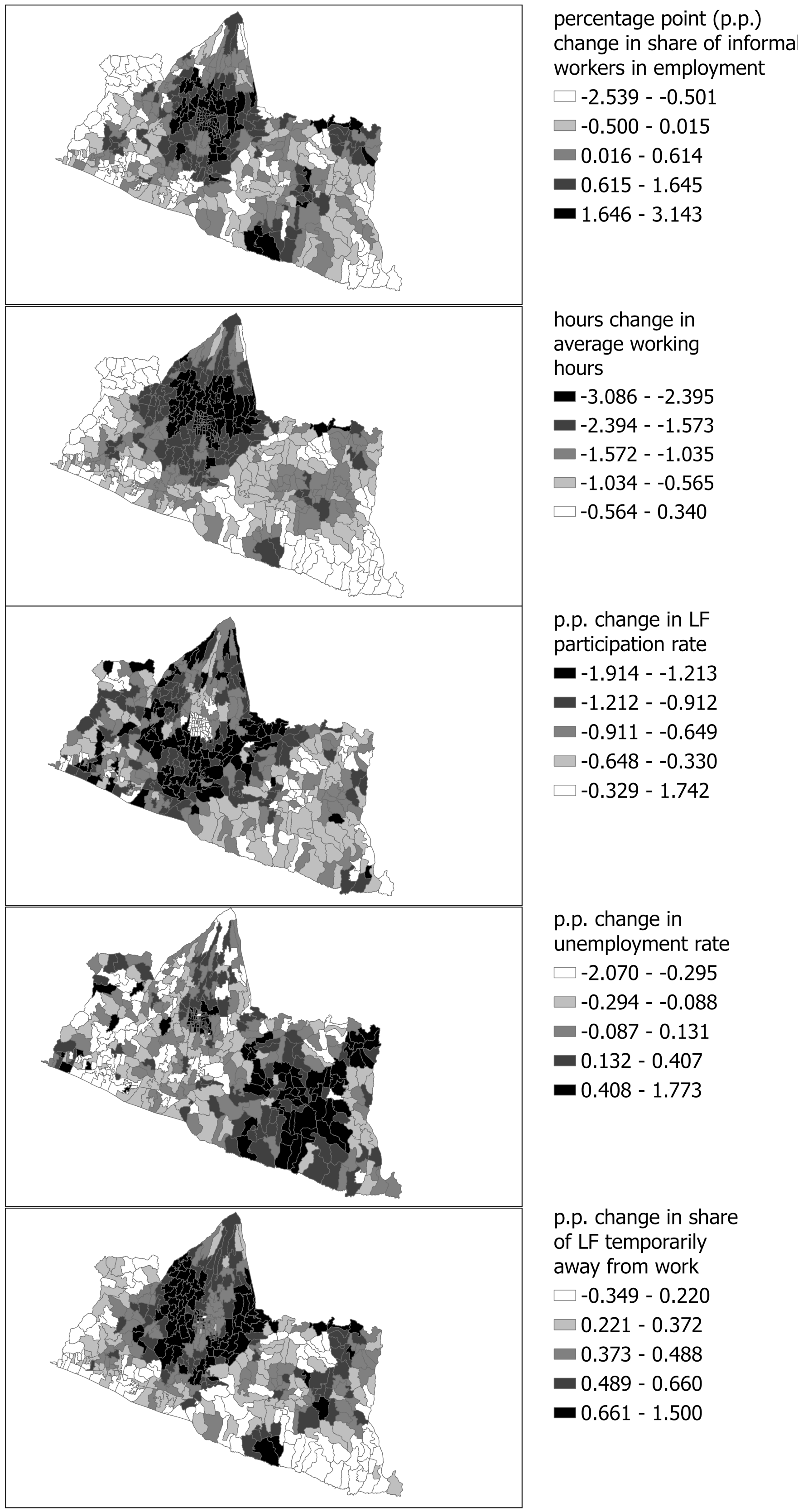

Figure 4: Changes in Labour Market Outcomes, 2019-2020

5.2 Estimation of Affected Workers

The estimation of the geographical distribution of the affected worker can be seen in Figure 4. Selected variables to examine are (i) the share of informal workers in employment, (ii) the labour force participation rate, (iii) the unemployment rate, and (iv) average working hours. The value of each variable is the difference between two static models: the model for the year 2020, which represents the post-COVID outbreak, and the 2019 for pre-outbreak.

Changes in the share of informal workers were concentrated in an area that spans from Kota Yogyakarta northward to Sleman. This pattern corresponds to the informal workers' distribution in 2019. One could suspect that this pattern signals the shift that happens for formal workers towards informal employment (Basri and Fitrania, 2022).

The drop in average working hours is concentrated in Kota Yogyakarta and Sleman. This pattern corresponds with the distribution in the change of informal employment share. A drop in working hours could indicate that shift to informal work is reducing the amount of work in general. Alas, correlation does not imply causation and this should be taken with a massive grain of salt. The first two indicators together might point to the vulnerability of formal workers, while early on, I point out that informal jobs are more exposed to negative consequences. This sounds like a contradiction, but Manning (2021) found that industries which consist of predominantly formal jobs are not less vulnerable than those with mostly informal jobs.

I will discuss the labour force participation rate and unemployment rate together. This is due to the empirical evidence that a significant portion of the affected workers left the labour force altogether instead of trying to find a job during the pandemic (see Section 2). The difference in geographic pattern is seen between the two variables. A drop in labour force participation is seen near Kota Yogyakarta, particularly its southern side in Bantul. While the rise in the unemployment rate happens within Kota Yogyakarta and in most parts of Gunung Kidul.

The share of the labour force temporarily stood down or away from work is also concentrated around the urban centre. The highest concentration forms a parallel north-south line on the eastern and western sides of the urban centre. This pattern is rather unique and does not correspond to the pattern of other variables.

Now that we see the results, I want to take the readers back to the discussion we had in Section 2. I argued that we lose important insights due to aggregation. I juxtaposed the change in working hours at the village and district level in Figure 5. We lose so many insights given the concentration of large working hours reduction spanning across Kota Yogyakarta up until the larger part of southern Sleman. We also see some concentrations of working hours loss in Gunung Kidul, but this won't appear if we only use the district-level data currently available.

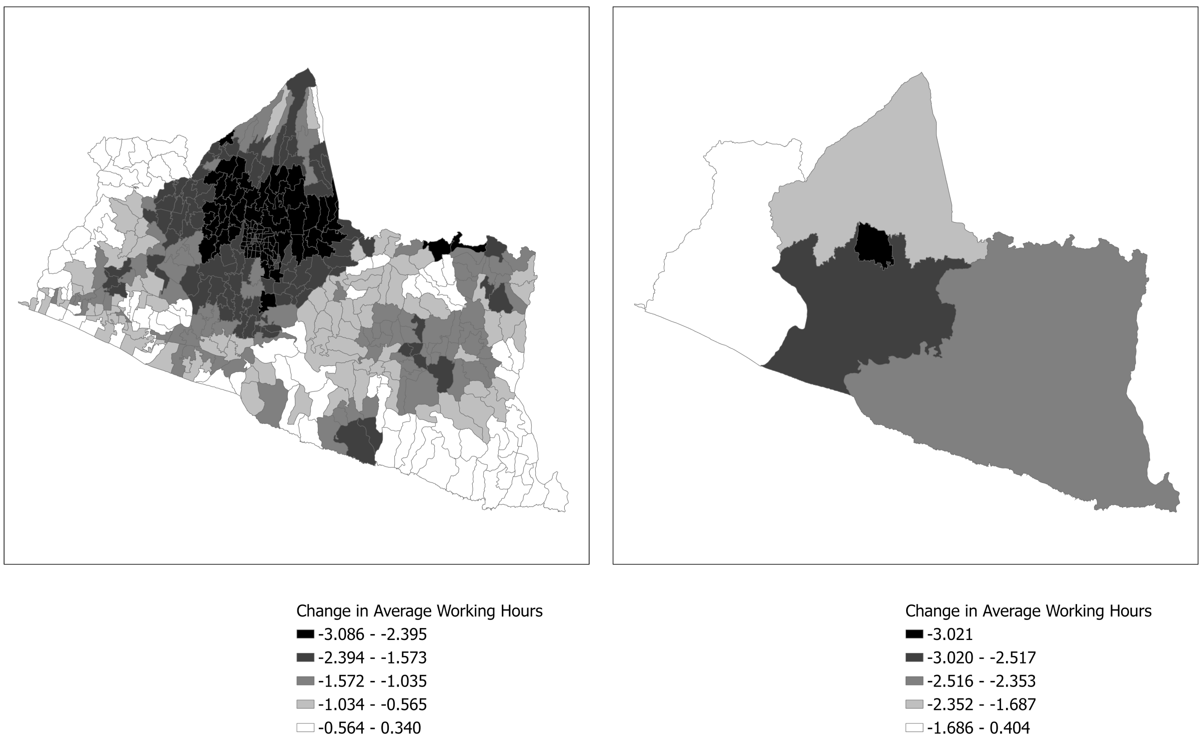

Figure 5: Comparison of Impact Mapping between Village- and District-level Data

5.3 What-if Analysis: Geographic Distribution of Hypothetical Income Loss

This section will demonstrate another potential of spatial microsimulation, the what-if analysis. The analysis is capable of mapping the distribution of an impact of a hypothetical scenario. In simple terms, we can answer the question 'How is the impact of a certain scenario distributed across space?' This is done by manipulating some of the variable values across all or part of the population microdata. The result of this change can then be seen through mapping in GIS. In addition to the mapping of the changes, I also performed Local Indicators of Spatial Association (LISA) (Anselin, 2010) and juxtapose it with the map of the impact to allow us a straightforward observation of the impact concentration.

I have defined two hypothetical scenarios based on general findings on COVID-19 impact on earnings. I attempt to demonstrate the method's capability to simulate changes in the population, both in whole and in part. The scenario can also be applied to a part of the population with particular attributes since we have the simulated population with all its employment characteristics.

The first scenario simulates the loss of income for the bottom half of employment in the earnings distribution in the accommodation, food and beverage sector. This represents how the most vulnerable workers in the tourism sector with the lowest productivity lost their jobs due to a sudden drop in tourism activity. The distribution of income loss can be seen in Figure 6. The loss of income is distributed mostly around Gunung Kidul in the eastern part of the province (shown in blue). LISA also shows that the concentrations of the most affected are also present in the western part of Sleman and Bantul near the urban centre.

The second scenario simulates the loss of income for all workers older than 50 presented in Figure 7. In this scenario, we see a large concentration of income loss around the dense part of the Sleman-Yogyakarta area (shown in blue). The large concentrations of the least affected are located around the coast in the southern part of Gunung Kidul (shown in red).

The impact of these scenarios is different for each geographical unit due to differences in the demographic composition of the population. This causes the spatial heterogeneity of the relationship between variables for each geographical unit. The what-if analysis of the spatial microsimulation has done well to illustrate this, as I have argued in Section 2.

This analysis can help better inform policymaking process by providing an ex-ante or ex-post analysis of a hypothetical scenario, including an assessment of a local-level policy impact or impact of a shock, as demonstrated in my case. The analysis helps by providing the geographical distribution of the impact. This section underlines further the core ability of the methods to estimate and identify geographical distributions of individual data to be used in various settings.

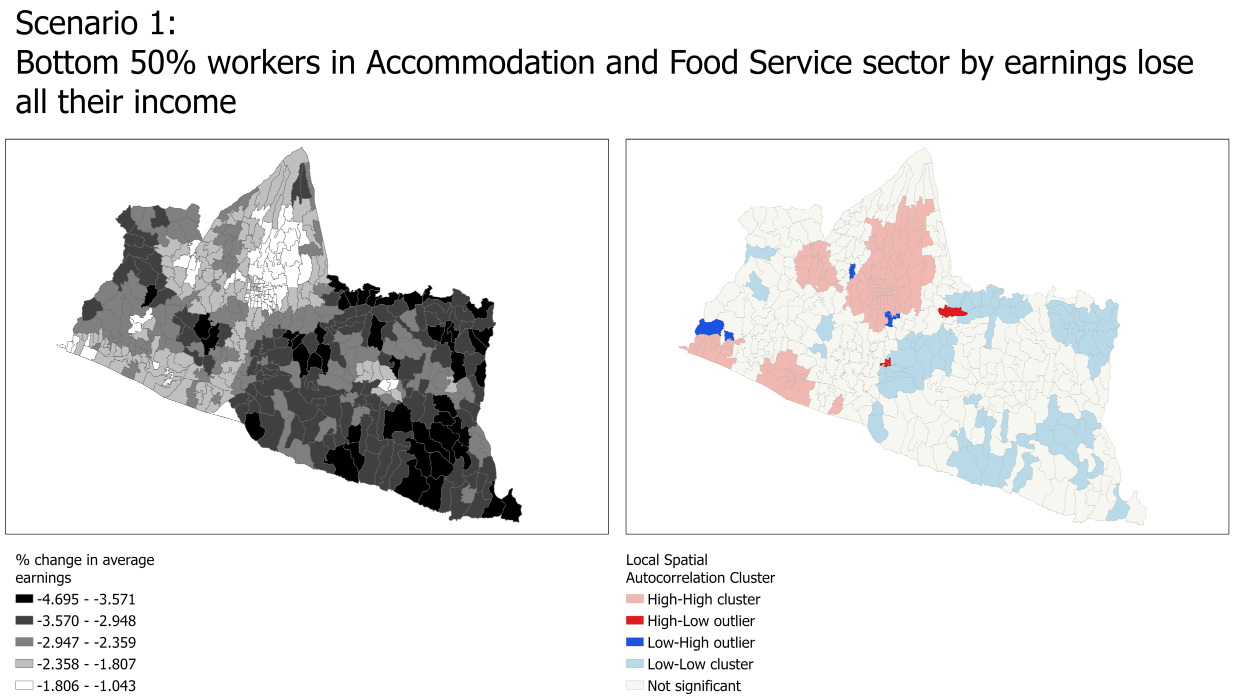

Figure 6: Scenario 1. Bottom 50% of Workers in Accomodation, Food, and Beverage Sector by Earnings Lose All of Their Income

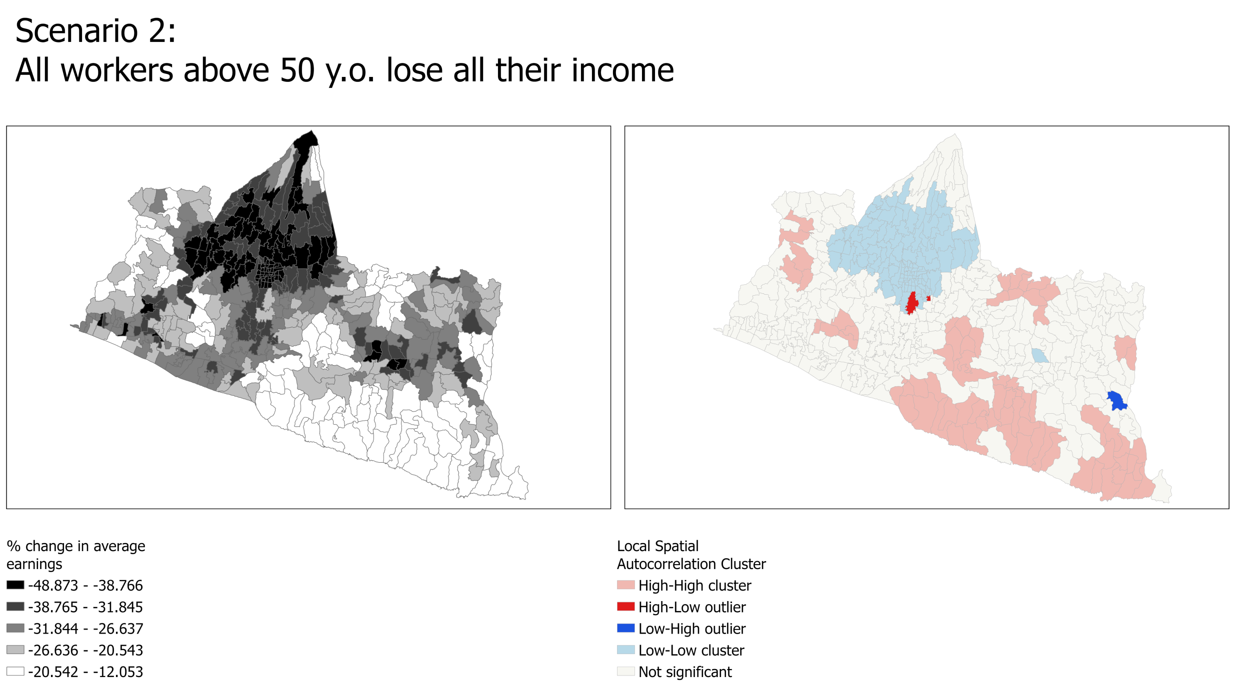

Figure 7: Scenario 2. All Workers Older Than 50 Lose All of Their Income

6 Concluding Remarks

The COVID-19 pandemic has overarching impact on our society, and we have yet to understand fully how it affects our livelihood. This paper raises concern about the role of geography and unveils its geographic pattern of economic impact, with a focus on labour market outcomes. I argue that previous studies focus too much on the 'who' rather than 'where', and even if 'where' is discussed, it severely lacks geographic insight.

I build *SimJogja*, the first spatial microsimulation model in Indonesia, and the first to use it to evaluate the geographic distribution of COVID-19 impact on the labour market outcomes in D.I. Yogyakarta. With this model, I demonstrate that spatial microsimulation, complemented with what-if analysis, can identify the spatial effects, namely spatial dependence and spatial heterogeneity, in the economic impact of COVID-19. This sets the stage for further investigation into the role of geography in affecting the economic outcomes of a negative shock.

I also show how spatial microsimulation can have policy relevance in this context. I show how the model could help COVID-19 assistance targeting by generating finely disaggregated geographic data. More importantly, I also demonstrate how ex-ante or ex-post assessment of scenario impact can be generated through what-if analysis to better inform policymaking processes.

Future development of the model is promising for a lot of uses. The model could be combined with *Survei Sosio-Ekonomi Nasional* (National Socio-economic Survey/SUSENAS) to estimate population microdata with socio-economic attributes. The model could also be combined with other methods to generate more insights from the estimation. One example by Ballas et al. (2006a) show how spatial microsimulation could be combined with data on commuting patterns and input-output tables to estimate the multiplier effect of a job loss. The model has great potential to further promote evidence-based policymaking practices in Indonesia.

7 References

Adams-Prassl, A., Boneva, T., Golin, M., Rauh, C., 2022. Work that can be done from home: Evidence on variation within and across occupations and industries. Labour Economics 74, 102083. https://doi.org/10.1016/j.labeco.2021.102083

Adams-Prassl, A., Boneva, T., Golin, M., Rauh, C., 2020. Inequality in the impact of the coronavirus shock: Evidence from real time surveys. Journal of Public Economics 189, 104245. https://doi.org/10.1016/j.jpubeco.2020.104245

Alon, T., Doepke, M., Olmstead-Rumsey, J., Tertilt, M., 2020. The Impact of COVID-19 on Gender Equality (No. w26947). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w26947

Andersson, F., Haltiwanger, J.C., Kutzbach, M.J., Pollakowski, H.O., Weinberg, D.H., 2018. Job Displacement and the Duration of Joblessness: The Role of Spatial Mismatch. The Review of Economics and Statistics 100, 203–218. https://doi.org/10.1162/REST_a_

00707
Angelucci, M., Angrisani, M., Bennett, D., Kapteyn, A., Schaner, S., 2020. Remote Work and the Heterogeneous Impact of COVID-19 on Employment and Health (No. w27749). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w27749
Anselin, L., 2010. Local Indicators of Spatial Association-LISA. Geographical Analysis 27, 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x
Aum, S., Lee, S.Y.(Tim)., Shin, Y., 2021. COVID-19 doesn't need lockdowns to destroy jobs: The effect of local outbreaks in Korea. Labour Economics 70, 101993. https://doi.org/10.1016/j.labeco.2021.101993
Baker, J.L., Grosh, M.E., 1994. Poverty reduction through geographic targeting: How well does it work? World Development 22, 983–995. https://doi.org/10.1016/0305-750X(94)90143-0
Ballas, D., Clarke, G., 2000. GIS and microsimulation for local labour market analysis. Computers, Environment and Urban Systems 24, 305–330. https://doi.org/10.1016/S0198-9715(99)00051-4
Ballas, D., Clarke, G., Dewhurst, J., 2006a. Modelling the Socio-economic Impacts of Major Job Loss or Gain at the Local Level: A Spatial Microsimulation Framework. Spatial Economic Analysis 1, 127–146. https://doi.org/10.1080/17421770600697729
Ballas, D., Clarke, G., Dorling, D., Rigby, J., Wheeler, B., 2006b. Using geographical information systems and spatial microsimulation for the analysis of health inequalities. Health Informatics Journal 12, 65–79. https://doi.org/10.1177/1460458206061217
Ballas, D., Rossiter, D., Thomas, B., Clarke, G., Dorling, D., 2005. Geography matters: Simulating the local impacts of national social policies, Contemporary research issues. Joseph Rowntree Foundation, York.
Basri, M.C., Fitrania, S.N., 2022. The Indonesian Economy in the time of COVID-19, in: COVID-19 in Indonesia. Routledge, London, pp. 12–45. https://doi.org/10.4324/9781003243670-2
Birkin, M., Wu, B., 2012. A Review of Microsimulation and Hybrid Agent-Based Approaches, in: Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M. (Eds.), Agent-Based Models of Geographical Systems. Springer Netherlands, Dordrecht, pp. 51–68. https://doi.org/10.1007/978-90-481-8927-4_3
Blocker, A.W., 2022. Ipfp: Fast Implementation of the Iterative Proportional Fitting Procedure in C.
Broomhead, T., Ballas, D., Baker, S., 2023. Oral health, sugary drink consumption and the soft drink industry levy: Using spatial microsimulation to understand tooth decay. Regional Science Policy & Practice rsp3.12682. https://doi.org/10.1111/rsp3.12682
Clarke, M., Forte, P., Spowage, M., Wilson, A.G., 1984. A Strategic Planning Simulation Model of a District Health Service System: The In-patient Component and Results, in: Davis, K., Van Eimeren, W., Van Eimeren, W., Engelbrecht, R., Flagle, C.D. (Eds.), Third International Conference on System Science in Health Care. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 949–954. https://doi.org/10.1007/978-3-642-69939-9_220
Coibion, O., Gorodnichenko, Y., Weber, M., 2020. Labor Markets During the COVID-19 Crisis: A Preliminary View (No. w27017). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w27017

Cowan, B., 2020. Short-run Effects of COVID-19 on U.S. Worker Transitions (No. w27315). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w27315

Dai, R., Mookherjee, D., Quan, Y., Zhang, X., 2021. Industrial clusters, networks and resilience to the Covid-19 shock in China. Journal of Economic Behavior & Organization 183, 433–455. https://doi.org/10.1016/j.jebo.2021.01.017

Delventhal, M.J., Kwon, E., Parkhomenko, A., 2022. JUE Insight: How do cities change when we work from home? Journal of Urban Economics 127, 103331. https://doi.org/10.1016/j.jue.2021.103331

Desmet, K., Wacziarg, R., 2022. JUE Insight: Understanding spatial variation in COVID-19 across the United States. Journal of Urban Economics 127, 103332. https://doi.org/10.1016/j.jue.2021.103332

Dingel, J.I., Neiman, B., 2020. How many jobs can be done at home? Journal of Public Economics 189, 104235. https://doi.org/10.1016/j.jpubeco.2020.104235

Edwards, K.L., Tanton, R., 2012. Validation of Spatial Microsimulation Models, in: Tanton, R., Edwards, K. (Eds.), Spatial Microsimulation: A Reference Guide for Users. Springer Netherlands, Dordrecht, pp. 249–258. https://doi.org/10.1007/978-94-007-4623-7_15

Elbers, C., Fujii, T., Lanjouw, P., Özler, B., Yin, W., 2007. Poverty alleviation through geographic targeting: How much does disaggregation help? Journal of Development Economics 83, 198–213. https://doi.org/10.1016/j.jdeveco.2006.02.001

Fairlie, R., Couch, K., Xu, H., 2020. The Impacts of COVID-19 on Minority Unemployment: First Evidence from April 2020 CPS Microdata (No. w27246). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w27246

González-Leonardo, M., Rowe, F., Vegas-Sánchez, A., 2023. A “donut effect”? Assessing housing transactions during COVID-19 across the Spanish rural–urban hierarchy. Regional Studies, Regional Science 10, 471–473. https://doi.org/10.1080/21681376.2023.2191684

Henderson, J.V., Turner, M.A., 2020. Urbanization in the Developing World: Too Early or Too Slow? Journal of Economic Perspectives 34, 150–173. https://doi.org/10.1257/jep.34.3.150

International Labour Organization, 2020. COVID-19 crisis and the informal economy: Immediate responses and policy challenges. International Labour Organization, Geneva.

Krugman, P., 1991. Increasing Returns and Economic Geography. Journal of Political Economy 99, 483–499. https://doi.org/10.1086/261763

Lee, S., Schmidt-Klau, D., Verick, S., 2020. The Labour Market Impacts of the COVID-19: A Global Perspective. The Indian Journal of Labour Economics 63, 11–15. https://doi.org/10.1007/s41027-020-00249-y

Lee, S.Y.(Tim)., Park, M., Shin, Y., 2021. Hit Harder, Recover Slower? Unequal Employment Effects of the Covid-19 Shock (No. w28354). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w28354

Lemieux, T., Milligan, K., Schirle, T., Skuterud, M., 2020. Initial Impacts of the COVID-19 Pandemic on the Canadian Labour Market. Canadian Public Policy 46, S55–S65. https://doi.org/10.3138/cpp.2020-049

Leyva, G., Urrutia, C., 2023. Informal labor markets in times of pandemic. Review of Economic Dynamics 47, 158–185. https://doi.org/10.1016/j.red.2022.01.002

Lomax, N., Norman, P., 2016. Estimating Population Attribute Values in a Table: “Get Me Started in” Iterative Proportional Fitting. The Professional Geographer 68, 451–461. https://doi.org/10.1080/00330124.2015.1099449

López-Bazo, E., Barrio, T.D., Artís, M., 2005. Geographical distribution of unemployment in Spain. Regional Studies 39, 305–318. https://doi.org/10.1080/00343400500087034

Lovelace, R., Dumont, M., 2016. Spatial microsimulation with R, The R series. CRC Press, Taylor & Francis Group, Boca Raton, FL.

Manning, C., 2021. 6 The labour market shock and policy responses to the coronavirus pandemic, in: Economic Dimensions of Covid-19 in Indonesia. ISEAS–Yusof Ishak Institute Singapore, pp. 94–113. https://doi.org/10.1355/9789814951463-009

Mion, G., Naticchioni, P., 2009. The spatial sorting and matching of skills and firms. Canadian Journal of Economics/Revue canadienne d'économique 42, 28–55. https://doi.org/10.1111/j.1540-5982.2008.01498.x

Orcutt, G.H., 1957. A New Type of Socio-Economic System. The Review of Economics and Statistics 39, 116. https://doi.org/10.2307/1928528

Panori, A., Ballas, D., Psycharis, Y., 2017. SimAthens: A spatial microsimulation approach to the estimation and analysis of small area income distributions and poverty rates in the city of Athens, Greece. Computers, Environment and Urban Systems 63, 15–25. https://doi.org/10.1016/j.compenvurbsys.2016.08.001

Pfeffermann, D., 2013. New Important Developments in Small Area Estimation. Statistical Science 28. https://doi.org/10.1214/12-STS395

R Core Team, 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramani, A., Bloom, N., 2021. The Donut Effect of Covid-19 on Cities (No. w28876). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w28876

Rephann, T.J., Makila, K., Holm, E., 2005. Microsimulation for Local Impact Analysis: An Application to Plant Shutdown*. Journal of Regional Science 45, 183–222. https://doi.org/10.1111/j.0022-4146.2005.00368.x

Rey, S.J., Montouri, B.D., 1999. US Regional Income Convergence: A Spatial Econometric Perspective. Regional Studies 33, 143–156. https://doi.org/10.1080/00343409950122945

Rodríguez-Pose, A., 2018. The revenge of the places that don't matter (and what to do about it). Cambridge Journal of Regions, Economy and Society 11, 189–209. https://doi.org/10.1093/cjres/rsx024

Singh, V., Shirazi, H., Turetken, J., 2022. COVID-19 and gender disparities: Labour market outcomes. Research in Economics 76, 206–217. https://doi.org/10.1016/j.rie.2022.07.011

Smith, D., Heppenstall, A.J., Campbell, M., 2021. Estimating Health over Space and Time: A Review of Spatial Microsimulation Applied to Public Health 4, 182–192. https://doi.org/10.3390/j4020015

Sun, Y.-Y., Sie, L., Faturay, F., Auwalin, I., Wang, J., 2021. Who are vulnerable in a tourism crisis? A tourism employment vulnerability analysis for the COVID-19 management. Journal of Hospitality and Tourism Management 49, 304–308. https://doi.org/10.1016/j.jhtm.2021.08.014

Tanton, R., 2013. A Review of Spatial Microsimulation Methods. The International Journal of Microsimulation 7, 4–25.

Tobler, W.R., 1970. A Computer Movie Simulating Urban Growth in the Detroit Region.

Economic Geography 46, 234. https://doi.org/10.2307/143141
Warr, P., Yusuf, A.A., 2021. Pandemic-induced De-urbanisation in Indonesia (No. 2020/08). Australian National University, Canberra.
Wilson, A.G., Pownall, C.E., 1976. A New Representation of the Urban System for Modelling and for the Study of Micro-Level Interdependence. Area 8, 246–254.
World Bank, 2020. Indonesia High-frequency Monitoring of COVID-19 Impacts Round 3.
Wren-Lewis, S., 2020. The economic effects of a pandemic, in: Economics in the Time of COVID-19. CEPR Press, London.
Zhao, L., Hessel, P., Simon Thomas, J., Beckfield, J., 2021. Inequality in Place: Effects of Exposure to Neighborhood-Level Economic Inequality on Mortality. Demography 58, 2041–2063. https://doi.org/10.1215/00703370-9463660

Appendices

List of Variables

Constraint Variables

Variables	Category
Age	1. 15 - 19 years old 2. 20 - 24 years old 3. 25 - 29 years old 4. 30 - 34 years old 5. 35 - 39 years old 6. 40 - 44 years old 7. 45 - 49 years old 8. 50 - 54 years old 9. 55 - 59 years old 10. 60 - 64 years old 11. 65 - 69 years old 12. 70 - 74 years old 13. 75 - 79 years old 14. 80 years old and older
Gender	1. Male 2. Female
Educational Attainment	1. Elementary School (SD) 2. Junior High School (SMP) 3. Senior High School (SMA) 4. Vocational School (Diploma) 5. University (Sarjana) 6. Didn't Graduate Elementary School (SD)
Occupational Category	1. Out of Labour Force 2. Unemployed 3. Military/Police Force 4. Public Service 5. Agricultural Worker 6. Entrepreneur/Business Owner 7. Formal Sector (e.g. office worker) 8. Other Occupations

Target Variables

- Working Hours
- Unemployment Rate
- Labour Force Participation Rate

- Share of Informal Worker
- Share of Employed Women
- Share of Employed Youth (15-29 years old)
- Sectoral Share of Employment

External Validation

Table 2: External Validation of Target Variables

Variable	2019			2020		
	Simulated	SAKERNAS	Difference	Simulated	SAKERNAS	Difference
Self-employed	17.02	17.86	-0.83	15.74	18.13	-2.39
Self-employed with temporary/family/unpaid workers	21.51	15.39	6.12	20.56	17.58	2.98
Self-employed with paid workers	2.83	4.04	-1.21	2.48	3.81	-1.33
Employee	44.26	43.77	0.49	42.26	38.78	3.48
Casual agricultural workers	4.86	1.97	2.88	4.49	1.70	2.78
Casual non-agricultural workers	7.32	5.12	2.20	8.63	5.52	3.11
Family/Unpaid Worker	2.20	11.84	-9.64	5.85	14.48	-8.63
Sectoral Share of Employment						
Agriculture	27.78	19.36	8.42	27.31	20.17	7.13
Mining	0.78	0.74	0.04	1.01	0.76	0.25
Manufacturing	14.93	17.03	-2.10	15.50	17.03	-1.53
Gas & Electricity	0.18	0.17	0.01	0.12	0.13	-0.01
Water & Waste	0.25	0.19	0.06	0.26	0.24	0.02
Construction	8.53	6.65	1.88	7.77	6.02	1.76
Wholesale Trade	14.74	19.03	-4.29	16.43	21.02	-4.59
Transport & Warehousing	2.90	3.95	-1.04	2.47	3.44	-0.97
Accomodation & FnB	7.78	9.39	-1.60	6.47	8.24	-1.77
Information & Communication	0.67	0.97	-0.30	0.96	1.42	-0.47
Financial Services	1.17	1.52	-0.34	1.00	1.36	-0.36
Real Estate	0.04	0.06	-0.03	0.11	0.16	-0.05
Business Services	1.74	2.42	-0.68	1.22	1.61	-0.39
Public Services	5.08	3.56	1.52	4.83	3.60	1.23
Educational Services	5.41	6.88	-1.47	4.98	6.15	-1.17
Social Services	1.80	2.15	-0.35	1.76	2.13	-0.36
Other Services	6.22	5.95	0.27	7.80	6.52	1.28
Working Hours	36.421	37.153	-0.73	34.796	35.340	-0.54
Unemployment Rate	5.54	3.04	2.50	5.59	5.38	0.21
Employment Rate	91.65	94.57	-2.92	91.10	91.38	-0.28

Temporarily Away From Work	2.81	2.39	0.41	3.31	3.24	0.07
Labour Force Participation Rate	69.67	72.83	-3.16	68.86	67.65	1.21
Share of Informal Workers	50.50	50.23	0.27	51.17	53.14	-1.97
Underemployment Rate	3.40	3.24	0.16	6.31	5.31	1.01
Share of Women in Employment	41.75	43.61	-1.86	43.18	44.20	-1.01
Share of Youth in Employment	18.41	22.07	-3.66	19.47	21.15	-1.68

Note: Every variables are in percentage, except for working hours which is in average hours