Modelling the Relationship of Infrastructure and Externalities Using Urban Scaling

Jung-Hun Yang and Kwang-Woo Nam *

Department of Urban Planning and Engineering, Kyungsung University, Busan 48434, Korea; junghun.yang1@gmail.com
* Correspondence: kwnam@ks.ac.kr; Tel.: +82-051-663-4794

Abstract: A surprising aspect of the agglomeration economy is the lack of attention to the impact on the physical environment. Even in the field of spatial planning, road infrastructure has been built in situations where the consideration of the agglomeration economy is insufficient. The urban scaling proposed by theoretical physicists is an excellent tool to solve this problem but is only at the level of conceptually comparing the index values extracted by individually scaling socioeconomic indicators and urban infrastructure with the population. Accordingly, the frame model scales the urban infrastructure with the number of workers by industry sector and includes a density externality structure so that the agglomeration economy and urban infrastructure can be linked directly. Three experiments were conducted to verify the frame model: first, the Zipf distribution of economic activity found straight lines in large cities, peaks in medium cities and hills in small cities; the cities were categorised by urban size. The second experiment verified that linearisation was due to Jacobs externalities, while the third confirmed that the peak was due to Marshall–Arrow–Romer externalities. Moreover, in distinguishing traditional and modern industries, thresholds of 0.6 in agglomeration and 1.0 in economic interaction were found.

Keywords: urban scaling; Zipf law; agglomeration economies; Jacobs externalities; MAR externalities; urban infrastructure

1. Introduction

Why do cities grow and decline? China faces this question, as its urban growth is steadily decreasing, and so does Korea, which has entered a phase of full-scale urban decline [1–3]. It is also an important issue in Western Europe and Japan, where urban decline has been progressing for a prolong time. Industrial structure policy is the most direct tool for city and national policymakers to deal with growth and decline [4]. This study explores urban growth issues related to the economy and industry.

Urban growth is closely related to the industrial structure. As the core theory, the agglomeration economy can be explained through externalities that occur when firms gather and interact in a specific region and are divided into static and dynamic externalities according to the time scale [5–7]. Marshall–Arrow–Romer (MAR) externalities induce knowledge spillover by geographically aggregating firms that belong to a specific industry. Representative areas include Silicon Valley [8] from the high-tech industry, Emilia-Romagna from small- and medium-sized traditional industries and Toyota City from the automobile industry.

The difference among these regions is that they are developing around high-tech venture firms, small and medium enterprises and large companies. A common aspect is that they are accumulated in small- and medium-sized cities [9,10]. Jacobs-type externalities refer to effects that stimulate innovation by the aggregation of several firms in various industries [11,12].

In other words, interactions among heterogeneous individuals accumulate over time and lead to innovation, which is the driving force for growth. Representative regions
include large cities, such as New York [13] and London. After the third industrial revolution, the benefits of agglomeration were strengthened across all industries [14]; however, two issues exist in relation to agglomeration economies. First, in terms of agglomeration economies, there is a surprising lack of interest in the impact of spatial structures, such as urban infrastructure. The connectivity among cities is significantly strengthened by infrastructure, such as road networks, which are estimated to work beyond the scope of small and large cities [15].

If a city is a product of socioeconomic dynamics formed by physical constraints [16], we can assume that its growth is determined by how MAR and Jacobs externalities overcome these constraints [15,17]. If the economic and industrial policy is a function-oriented approach that deals with urban growth, a mathematical basic frame is needed as a starting point to link functions and forms (i.e., to converge economic policies with spatial planning). Van Soest, Gerking and van Oort [18] argued that the spatial effect of the agglomeration economy has important implications for the development of endogenous growth models and urban planning policymaking. In this aspect, if the mathematical basic frame of this study can provide parameters that can conceptually control the endogenous growth model and urban planning, it could play an important role in public policy dealing with the economic function and spatial form of the city.

Second, the scope of the agglomeration economy is believed to be restricted mainly to cities; however, there is no consensus on what a city is and how it should be described [19]. Particularly, the traditional segregation between city centres and suburbs in large cities of the twenty-first century is being dissolved, and cities are being transformed into vast interconnected areas that merge with other cities [20]. This phenomenon suggests that MAR externalities of small cities may disseminate the effects of agglomeration economies [19]; simultaneously, Jacobs externalities for large cities may function throughout the convergence of several metropolitan areas.

Furthermore, the measurement result of the agglomeration economy may vary depending on the Modifiable Areal Unit Problem (MAUP). This means that the results of the study may vary depending on the spatial unit selected [21]. Thus, measurement of the agglomeration economy could be affected by the MAUP [19,22–24]. However, Amrhein and Flowerdew [25] found that there was no significant difference in the mean values of the measurements performed repeatedly with various compartmentalization methods. Hence, they could not find any evidence for the compartmental effect. As a leading study, Van Soest et al. [18] found that the agglomeration economy works well on a geographical scale smaller than that of a city. However, Andersson, Klaesson and Larsson [26] indicated that MAUP has often been discussed in the field of geography dealing with space, however, less often in the field of economics, noting that the economics approach avoids the problem of shape and size [27–29].

As a solution, it is necessary to consider the expansion of the spatial range of externalities to the mega-region [30]; therefore, in this study, experiments are conducted in an area that is presumed to be a mega-region. In addition, it is strategically structured so that the measurement model of the agglomeration economy can have meaning according to the spatial scope. The spatial scope of interest is the mega-region and city.

The facts discovered through the experiment reflecting the solutions are as follows. We first developed a frame model in which economies of scale that occur in urban infrastructure and aggregation economies that occur in socioeconomic interactions can be traded off with each other. Thus, it is possible to quantitatively understand how urban infrastructure spatial planning affects each manufacturing sector and how economic policies, such as industrial clusters, can present an appropriate level of urban infrastructure.

Second, we verified that the slope of the Zipf distribution is the level of interaction between individuals. The shape of the slope is divided into straight lines, peaks and hills for each urban scale section. Dynamic externality models of aggregation, which are extensions of the frame model, were developed and explained for each. When the type of integrated externality changes, the gradient’s shape also changes. This shift suggests that it can act as
a basic structure that can control spatial planning and economic policy simultaneously in space and time.

Third, we found a critical point to distinguish between traditional and modern industries in the parameter phase space between the Marshall index and the coefficient. The Marshall coefficient critical point for economies of scale is 0.6, and the Marshall coefficient critical point for socioeconomic interaction is 1.0. They provide a measure of whether industrial sectors are efficient, in terms of economics and temporal and spatial distribution.

The research topic mentioned above is ‘Why does a city grow and decline?’ Through the literature review section, this study examines the dynamic externalities of agglomeration, such as MAR and Jacobs externalities, as a driving force for growth. Two aspects were derived as limits of the agglomeration externalities: the linkage with urban infrastructure and the scope of the agglomeration. The Zipf law and urban scaling have been reviewed as models to develop a frame, integrating economic and spatial factors. In the Materials and Methods section, mathematical integrated frame models are presented and the key models are composed, as follows:

- Socioeconomic interaction coefficient model.
- Jacobs model.
- Marshall model.

Further, in the Research Methods section, model description and study design are presented to examine whether the proposed frame model verifies the three research hypotheses. In the Results and Discussion sections, the results of testing the three experimental hypotheses are analysed. This study ends with the summary and suggestions for further research.

2. Literature Review

Rybski, Arcaute and Batty [31] asked ‘How does the form of a city determine their function?’ and indicated location theory [32,33], bottom-up approach [34,35], scaling [36,37] and the Zipf law [38,39] as related studies. To link the functions and forms mentioned above, it is necessary to minimise the influence of the environmental or individual factors of industrial functions and spatial structures. Therefore, a methodology with properties that transcend contextual constraints is required, which includes the fields of scaling and the Zipf law [16,39,40]. However, as the Zipf law is a special form of scaling [41], scaling eventually becomes the focus.

Zipf [39] argued that people behave in a way that requires the least effort and showed that the probability distribution of city rank according to population size represents a power law of a linear graph with a slope of $-1$. Urban geographers and economists argue that these rank-size rules reflect the hierarchy of centres [42]. Based on this logic, another way of expressing the fact that the relation between rank and size is linear is that the distribution of city sizes adheres to the power law [43]. As a typical form of the Zipf law, Equation (1) indicates the population of the city ranked $r$:

$$P_r = P_1 r^{-\alpha},$$

where $r$ is the city rank, and $\alpha$ is a parameter controlling the population differences among cities. For a typical $\alpha = 1$, the population of the second city is half of the population of the first city, and that of the third city is one third. As the value of $\alpha$ increases (i.e., as the slope of the probability distribution becomes steeper), it indicates polarisation where most of the population is concentrated in the first-ranked city. Zipf [38,39] argued that the linear distribution of the $-1$ slope implies that people behave as groups not as individuals. Bettencourt and his colleagues in the field of statistical physics, found that socioeconomic output changes superlinearly and physical volume changes sublinearly.
as the city population increases [36,37]. This is called the urban scaling theory and is commonly characterised by a power function [36,37,44,45], as demonstrated:

\[ Y = \alpha X^\beta \]  

(2)

where \( Y \) is the urban indicator of a socioeconomic quantity, such as the number of professional workers, or a physical quantity, such as the number of gas stations, and \( X \) is the size of a city generally expressed in terms of city population. \( \alpha \) is the constant of proportionality, and \( \beta \) is the scaling exponent.

When socioeconomic indicators are applied to \( Y \), the scaling index has a value greater than 1, indicating a superlinear relationship, in which socioeconomic quantities increase more rapidly and at a larger proportion compared to the population. This implies an increasing return to scale. Conversely, when physical indicators, such as urban infrastructure are applied, the scaling index has a value less than 1, indicating a sublinear relationship in which the physical quantity increases slowly compared to the population growth. This implies scale economies [37,46].

A linear system with a scaling index close to 1 is an indicator related to individual needs, which appears when housing and household electricity consumption are applied [16]. West [46] obtained values of 1.15 for the superlinear index and 0.85 for the sublinear index through empirical analysis and estimated that the physical and the social quantities are mirror images of each other. In other words, the reduced socioeconomic activity energy secured through economies of scale in urban infrastructure can be a driving force to increase the returns of social interaction by 15%.

The issues raised in relation to scaling are as follows. First, population size alone does not provide sufficient information to explain the state of a city. Most city indicators extend linearly with city size; however, if there is a non-linear relationship, the index fluctuates considerably [16]. A systematic explanation of the process that causes nonlinear scaling is required [31]. Second, it is necessary to understand how the structure of increasingly complex socioeconomic organisations, such as enterprises and cities, changes according to size [46]. How this change in the scaling index is related to the evolution of socioeconomic organisations and their principles are not well understood.

3. Research Methods

3.1. Model Description

As the first step in linking the function and shape of a city, a basic mathematical model is presented using the standard urban scaling function [36,37], which is shown in Equation (2). The function–shape model is the scaling between the number of employees and the length of the road, shown as:

\[ L = \alpha S^\beta \]  

(3)

where \( L \) is the road extension, and \( S \) is the size of the city as the total number of workers employed by the firm within the city. The parameter \( \alpha \) is a coefficient representing the relationship between \( S \) and \( L \), while \( \beta \) is a scaling exponent. Equation (3) shows how spatial system \( L \) reacts when industrial size \( S \) changes. Applying socioeconomic indicators (such as GDP) to \( L \) will show superlinear scaling, with \( \beta > 1 \), thus, indicating increasing returns. However, this study is focused on the relationship with urban infrastructure. When \( \beta < 1 \), it is called sublinear scaling, which demonstrates economies of scale where infrastructure investment costs decrease systematically as the number of workers increase.

Next, this study develops a frame model by adding the density concept to the basic model. The spatial density of economic activity is also known as density externality [26]. The core frame model of this study is the addition of the concept of density externality to the basic model of Equation (3), as follows:

\[ \left( \frac{L}{A} \right) = \alpha \left( \frac{S}{A} \right)^\beta \]  

(4)
where A represents the area. Equation (4) shows how the road extension varies, depending on the number of employees and how the efficiency of urban infrastructure plays a role in the dynamics of the industry. The biggest difference between the basic and frame models is that the proximity among individuals is controlled by the density and diversity of the industry size [47].

The formula of the frame model contains the integration externality of the scaling structure and the density externality of the variable itself. It can be interpreted as a combination of scaling externalities and density externalities. In addition, the density structure has the advantage of controlling spatial heterogeneity between cities. However, the concept of spatial density reveals little about the geographic scale involved. The spatial scale and density of the analysis is a matter of assumptions about the scale at which externalities operate [26].

As a solution direction, this study focuses on the attribute composition of density, rather than the structural change of the frame model. Externality operating across all industrial entities affects the overall spatial scale of a wide area beyond the city scale. It is estimated that externality operating by the industry sector affects the economic activities within the city scale [17]. This step can be summarized with two assumptions:

**Hypothesis 1 (H1).** The scale of the entire industry is important for externalities operating at the regional level.

**Hypothesis 2 (H2).** The scale of each industry is important for externalities operating at the city level.

Each of these hypotheses is linked to the Jacobs model in Equation (5) and the Marshall model in Equation (6).

In line with the dynamic externalities reviewed in the previous section, this study integrates Jacobs and MAR externalities with the infrastructure problem using Equation (4), the frame model. First, Jacobs externalities promote innovation by accumulating interactions caused by aggregation of heterogeneous sectors over time. The size S of the frame model is the number of workers employed by firms located in the city in Equation (5) as the Jacobs model,

\[
\left( \frac{L}{A} \right) = \alpha_j \left( \sum S_d \frac{S_d}{A} \right)^{\beta_j},
\]

where \( L \) is the road extension, \( A \) is the city area, and \( S_d \) is the number of employees in the division field. While \( \alpha_j \) is the Jacobs coefficient, \( \beta_j \) is the Jacobs exponent.

The Jacobs coefficient \( \alpha_j \), which represents the level of interaction among individuals, shows how well the socioeconomic dynamics, represented by the diversity and size of industries, can utilise urban infrastructure as a road network. The Jacobs index \( \beta_j \) represents the level of the scale economies of urban infrastructure compared with the entire industry.

Next, MAR externalities cause knowledge diffusion by accumulating homogeneous sectors. Equation (6), as the MAR model, indicates Marshall coefficients and exponents for each division, as follows:

\[
\left( \frac{L}{A} \right) = \alpha_{md} \left( \frac{S_d}{A} \right)^{\beta_{md}},
\]

where \( L \) is the road extension, \( A \) is the city area, and \( S_d \) is the number of employees in the division field. \( \alpha_{md} \) is the Marshall coefficient, and \( \beta_{md} \) is the Marshall exponent. As the coefficients and exponents of the MAR model are calculated for each division, these values are expected to show a specific pattern in the parameter space, where the x and y axes represent the exponents and coefficients, respectively.
The following model is the interaction coefficient representing the socioeconomic status of the region:

$$a = \left( \frac{L}{A} \right) / \left( \frac{S}{A} \right)^\beta$$  \hspace{1cm} (7)

The denominator is the density of workers, which increases or decreases in a power form. The numerator is the density of the road network. The usable road network density can be used per unit density of workers. As the road network enhances and promotes social interaction, the coefficient value increases as the road network density increases.

3.2. Study Design

One frame model and two externality models are tested in this study. Thereafter, the behaviour of socioeconomic interaction coefficients by industry division is observed. The three models are linked to the research hypotheses mentioned in the introduction and are verified by three experiments.

To solve the two issues mentioned above, this study presents basic ideas and experimental hypotheses for hypothesis establishment. First, the basic idea is that Jacobs externality has a global effect and Marshall externality has a local effect [17]. In other words, Jacobs externality, which operates over a megaregion beyond a city or metropolitan area, gives order to the systems of cities, which is assumedly represented by the slope of the Zipf distribution of $-1$. On the contrary, Marshall’s externality cannot easily exceed the spatial range of individual cities, which is presumed to affect the irregular slope. The experimental hypotheses for the research design are as follows:

**Hypothesis 3 (H3).** The Zipf distribution in economic and spatial agglomerations shows regular linear lines and irregular peaks and hills.

**Hypothesis 4 (H4).** Regular linear lines are expressed by Jacobs externalities.

**Hypothesis 5 (H5).** Irregular peaks and hills are expressed by Marshall externalities.

Experiment 1 shows the distribution of city rank and employment size as Zipf model scaling. We examine whether the Zipf distribution shows regular straight lines and irregular hills. Jacobs externalities lead to innovation by accumulating interactions generated by the large-scale aggregation of heterogeneous sectors over time, eventually leading to urban growth [48–50]. In this regard, the hypothesis of this study is that the regular linear distribution of the Zipf law is expressed by Jacobs externalities. We analyse whether this hypothesis is verified in Experiment 2.

Experiment 2 is the Jacobs model scaling, which tests the hypothesis that Jacobs externalities are the driving force behind the straight-line pattern of the Zipf plot. In test step 1 of Experiment 2, feature scaling is performed to adjust the size and range to include the influence of variables evenly. After log transformation, the range of the variable values is converted into (0,1) by normalisation. Equation (8) is the min–max normalisation. The scaled value is given by:

$$v_{scaled} = \frac{v - v_{min}}{v_{max} - v_{min}}$$  \hspace{1cm} (8)

where $v$ is the variable value, $v_{min}$ is the minimum value in the variable list, and $v_{max}$ is the maximum value in the variable list. Test step 2 is a regression analysis. By taking the logarithm of both sides of Equation (6), a regression equation is derived, as shown in Equation (9). This is given by:

$$\ln \left( \frac{L}{A} \right)_i = \ln a_j + \beta_j \ln \left( \frac{\sum d}{A} \right)_i$$  \hspace{1cm} (9)
Experiment 2 performs regression analysis of the linear equation obtained by the log transformation of Equation (6), which is the power equation of employment density for road density.

The hypothesis of the third experiment is that irregular hill scaling in the urban system of the Zipf model is expressed by Marshall externalities. Rosenthal and Strange [51] argued that agglomeration economies tend to be prominent in the localisation economy, which are the economies of scale that occur within a specific industry. Sarkar et al. [45] estimated that the localisation economy within a specific industry shows superlinearity in scaling, by applying socioeconomic indicators. Hence, in the urban system of Experiment 1, the peaks of medium cities and the hills of small cities represent regional aggregation; our study suggests that this is due to MAR externalities. Experiment 3 is the Marshall model scaling. The Marshall exponents and coefficients are obtained for each of the 19 manufacturing divisions, shown as:

\[ \ln \left( \frac{L_i}{A_i} \right) = \ln \alpha_{md} + \beta_{md} \ln \left( \frac{S_d}{A_i} \right) \]  (10)

The pattern in which 19 Marshall exponents and coefficient pairs are distributed in the parameter phase space is identified, and the thresholds are analysed. The time range is applied differently in each experiment. Experiment 1 is interested in discovering a pattern that gradually changes annually over a long period of time, and thus we trace the long-term growth fluctuations over a 24-year period, from 1994 to 2017.

As Masan and Jinhae were independent cities for 18 years before they were integrated into Changwon in 2010, for consistency of analysis, the integrated Changwon region is divided into Changwon, Masan and Jinhae districts. Therefore, the total number of target city groups in Experiment 1 is 22. As Experiment 2 and Experiment 3 are interested in comparing changes between specific time periods, Changwon is integrated into one area for these experiments. The total number of target cities and counties is 20. Experiment 2 applies to 2006 and 2017; Experiment 3 applies to 2006, 2011 and 2017.

4. Study Area and Dataset

As this study focuses on why cities grow and decline, the south-eastern region of Korea is an interesting area for research (Figure 1). The emergence, growth and decline of the manufacturing industry in the south-eastern region have progressed rapidly in a short time span. This area has been the central region of the Korean manufacturing industry, focusing on the light industry in the 1960s, the heavy chemical industry in the 1970s and the automobile, ship and mechanical parts industries since the 1980s.

However, post 2000, the growth in the south-eastern region has weakened; the Korean government is promoting growth in this region through economic and industrial policies, centred on next-generation industries, such as electric vehicles and spatial planning policies, such as the construction of a wide area railway network, road network and new airport. Particularly, local governments in the south-east region are strategically attempting a mega-region that includes all cities and counties in the area [52,53].

The dataset consists of the number of employees in firms and the length of the road network. The number of employees in individual firms is based on the ‘national business survey’ data, provided by the microdata service system of the Korea National Statistical Office (http://mdis.kostat.go.kr accessed on 1 March 2020). Microdata are obtained by removing input errors from raw data and are used as basic data for data processing, such as creating statistical tables. Each record consists of the address of an individual firm (city and district code), industry sector (single letter and two and three digit codes) and number of employees. Next, the employment size of individual firms in each city and county is determined based on the division (two digits) of the manufacturing section (single letter) in the Korean Standard Industrial Classification (KSIC) from the national business survey data (http://www.kosis.kr accessed on 1 March 2020).
For road data, the current status data by road type are used, as provided by the National Statistics Portal (http://www.kosis.kr accessed on 1 March 2020). There are five types of roads in Korea, and each has a different transport capacity. National expressways are automobile-only roads that connect cities that are central to the transportation network; general national highways are national key road networks that connect major cities, designated ports and major airports; local highways are a network of local highways connecting metropolitan cities and counties; metropolitan city roads are roads within a metropolitan city; and City and county roads are roads within general cities and counties. As this study considers both intra- and inter-city interactions, it is summed up by applying weights for each road type, as shown in Equation (11). The road extension $L$ is shown as:

$$L = (8 \cdot l_1) + (2 \cdot l_2) + (1 \cdot l_3) + (0.5 \cdot l_4) + (0.5 \cdot l_5)$$  \hspace{1cm} (11)$$

where $l_1$ is a national expressway, $l_2$ is a general national highway, $l_3$ is a local highway, $l_4$ is a metropolitan city road, and $l_5$ is a city and county road.

Data pre-processing is the most time-consuming process in the field of data science dealing with big data. There are three major data corrections in this study, which are not error correction of raw data but the rearrangement of changes in general administrative districts and industry classification codes based on the previous year. First, by correcting the administrative district map for 23 years (between 1994 and 2016) on the GIS based on the administrative district map for 2017, the city codes of individual companies for 23 years can be organized as of 2017.

Second, the major and medium classification codes of individual firms are revised based on the 2017 KSIC standards for the industry classification standards, which were changed between 1994 and 2017. The code conversion table is provided by Statistics Korea (http://kostat.go.kr accessed on 1 March 2020).

Figure 1. Study area in south-eastern Korea.
Third, six divisions are excluded from the 25 divisions in KSIC’s manufacturing middle class (two digits). If firms are located in only a few cities and counties, the analysis results may be distorted. The manufacturing divisions (two digits) that were not properly distributed in the 20 cities and counties of Gyeonggi Province are excluded.

5. Results
5.1. Zipf Scaling

The key question of this study is why cities grow and decline. One must first determine whether a city is actually growing or declining. The Zipf distributions show how growth and decline appear in urban systems. First, this study examines the characteristics of each period. As shown in Figure 2, the 22 cities are arranged with slight deviations on a straight line with negative slope. The line in 1994 is close to a relatively smooth line. For about 10 years, between 1994 and 2003, the industrial size of intermediate cities increased significantly. Industrial growth in large cities decreased for approximately 14 years, between 2004 and 2017.

As a characteristic of each size, the linearity of the first to third largest cities, Busan, Ulsan and Changwon, becomes closer to a straight line over time. The linearity of large cities sets the standard for the growth of the entire city system. For example, if the size of employment is located above the line of a large city, we can presume it to be a city with centripetal force in the urban system. The linearity of the middle cities, ranked fourth to eighth, shows peaks at two points on the graph.

For the industry size of Gimhae, ranked fourth, the first peak rapidly approaches that of Changwon, which is ranked at the top. For Goeje’s industrial scale, ranked eighth, the second peak becomes similar to that of Yangsan, ranked fifth. In 2017, the industrial sizes of the fifth to eighth largest cities are similar. Changes in economic size among middle cities strengthen the irregularity of urban systems and changes in industrial growth exceed the alignment L2 of urban systems in large cities. The industrial growth of small cities, ranked 9th to 22nd, is lower than that of large cities.

As a main result, the urban system linearity responds differently depending on the size of the city. Large cities show strong regularity, middle cities show large variations in irregularities, and small cities show small variations in irregularities.
5.2. Jacobs Scaling

The coefficient of determination is 0.7981 in 2006 and 0.8477 in 2017, and the regression line in the upper left direction shown in Figure 3 has explanatory power. The scaling exponent $\beta$ increases slightly to 0.8174 in 2006 and 0.8296 in 2017, showing that economies of scale in the road network have weakened. The Jacobs externality coefficient $\alpha$ is 1.0696 in 2006 and decreases slightly to 1.0483 in 2017. It can be observed that the overall interaction of the manufacturing sectors has weakened.

![Figure 3. Jacobs scaling (Left: 2006, Right: 2017).](image)

As shown in Figure 3, the distribution of cities in 2006 and 2017 is well arranged around the regression line. However, Geoje is a singularity with a large deviation from the regression line. As shown in Figure 1, Geoje is presumed to be an island. The lower-right part of the graph (Figure 3) is where economies of scale exist and the upper-left part is the area of diseconomies of scale. For example, Changwon, located at points (0.8, 0.6) in 2006, requires only 0.75 units of road density when the employment density increases by one unit. Haman, located at points (0.35 and 0.5), requires 1.43 units of road density when the employment density increases by 1 unit. Therefore, we can assume that Changwon has a significantly higher growth potential than Haman in terms of Jacobs externalities, which reflect the scale economies of urban infrastructure and the strength of socioeconomic interaction.

By observing the city’s changes between 2006 and 2017, the cities that moved toward scale economies are Changwon, Gimhae, Yangsan, Sacheon and Haman; the cities that moved toward scale diseconomies are Ulsan, Geoje and Tongyeong. The cities with slight movement are Busan and Jinju.

5.3. Marshall–Arrow–Romer Scaling

Tables 1–3 summarises the Marshall exponents and coefficients estimated for each industrial division using regression Equation (10). The divisions belonging to the range of determination coefficients, and the significance by year is as follows. The main features can be divided into two categories: First, as time passes, the number of significant sectors decreases. Second, while traditional industries tend to have different significant sectors by year, modern industries tend to maintain the same sector.
### Table 1. Summary of exponents and coefficients (2006).

| Type               | Division                                      | Name                                      | 2006               |        |        |        | adj. $R^2$ |
|--------------------|-----------------------------------------------|-------------------------------------------|--------------------|--------|--------|--------|-----------|
|                    |                                               |                                           | $\beta$            | $p > |t|1$ | $\ln \alpha$ | $p > |t|1$ | $\alpha$ |          |
| Traditional Industry |                                               |                                           | 0.712              | 0.000  | 0.029  | 0.614 | 0.971 | 0.637       |
|                    |                                               | Textiles except apparel                   | 0.599              | 0.000  | 0.040  | 0.546 | 0.961 | 0.579       |
|                    |                                               | Wearing apparel, clothing                | 0.849              | 0.000  | 0.104  | 0.134 | 0.901 | 0.648       |
|                    |                                               | Wood and products of wood                | 0.719              | 0.000  | 0.120  | 0.075 | 0.886 | 0.685       |
|                    |                                               | Pulp, paper and paper products           | 0.383              | 0.002  | 0.108  | 0.070 | 1.115 | 0.392       |
|                    |                                               | Printing and reproduction                | 0.886              | 0.000  | 0.182  | 0.011 | 0.833 | 0.736       |
|                    |                                               | Rubber and plastic products              | 0.436              | 0.001  | 0.073  | 0.211 | 1.076 | 0.469       |
|                    |                                               | Other non-metallic minerals              | 0.648              | 0.004  | 0.126  | 0.311 | 0.882 | 0.339       |
|                    |                                               | Furniture                                | 0.592              | 0.000  | 0.022  | 0.707 | 0.978 | 0.623       |
| Modern Industry    |                                               | Basic metals                             | 0.412              | 0.000  | 0.106  | 0.050 | 1.112 | 0.482       |
|                    |                                               | Fabricated metal products                | 0.618              | 0.000  | 0.082  | 0.263 | 0.921 | 0.585       |
|                    |                                               | Electronic components                    | 0.491              | 0.000  | 0.117  | 0.008 | 1.124 | 0.635       |
|                    |                                               | Chemicals and chemical                   | 0.514              | 0.000  | 0.084  | 0.070 | 1.088 | 0.623       |
|                    |                                               | Medical, precision and optical           | 0.498              | 0.000  | 0.111  | 0.019 | 1.117 | 0.590       |
|                    |                                               | Electrical equipment                     | 0.550              | 0.000  | 0.200  | 0.707 | 1.020 | 0.622       |
|                    |                                               | Other machinery and equipment            | 0.457              | 0.000  | 0.039  | 0.475 | 1.040 | 0.569       |
|                    |                                               | Motor vehicles, trailers                 | 0.376              | 0.001  | 0.073  | 0.247 | 1.076 | 0.414       |
|                    |                                               | Other transport equipment                | 0.407              | 0.002  | 0.066  | 0.309 | 1.071 | 0.379       |
| others             |                                               | Other manufacturing                     | 0.859              | 0.000  | 0.189  | 0.026 | 0.828 | 0.654       |

Cell color: $p < 0.1$ and adjusted-$R^2 > 0.5$. Division full name: (1) Food products; (2) textiles except apparel; (3) wearing apparel, clothing accessories and fur articles; (4) wood and products of wood and cork, except furniture; (5) pulp, paper and paper products; (6) printing and reproduction of recorded media; (7) rubber and plastics products; (8) other non-metallic mineral products; (9) furniture; (10) basic metals; (11) fabricated metal products, except machinery and furniture; (12) electronic components, computers, and visual, sounding and communication equipment; (13) chemical and chemical products, except pharmaceuticals and medicinal chemicals; (14) medical, precision and optical instruments, watches and clocks; (15) electrical equipment; (16) other machinery and equipment; (17) motor vehicles, trailers and semitrailers; (18) other transport equipment; and (19) other manufacturing (jewellery and related articles, musical instruments, sports, athletic goods, etc.).

### Table 2. Summary of exponents and coefficients (2011).

| Type               | Division                                      | Name                                      | 2011               |        |        |        | adj. $R^2$ |
|--------------------|-----------------------------------------------|-------------------------------------------|--------------------|--------|--------|--------|-----------|
|                    |                                               |                                           | $\beta$            | $p > |t|1$ | $\ln \alpha$ | $p > |t|1$ | $\alpha$ |          |
| Traditional Industry |                                               |                                           | 0.685              | 0.000  | 0.057  | 0.419 | 0.945 | 0.563       |
|                    |                                               | Textiles except apparel                   | 0.525              | 0.000  | 0.010  | 0.884 | 0.990 | 0.477       |
|                    |                                               | Wearing apparel, clothing                | 0.726              | 0.000  | 0.015  | 0.809 | 0.985 | 0.559       |
|                    |                                               | Wood and products of wood                | 0.534              | 0.000  | 0.012  | 0.856 | 1.011 | 0.572       |
|                    |                                               | Pulp, paper and paper products           | 0.388              | 0.003  | 0.090  | 0.159 | 1.094 | 0.563       |
|                    |                                               | Printing and reproduction                | 0.808              | 0.000  | 0.176  | 0.068 | 0.881 | 0.674       |
|                    |                                               | Rubber and plastic products              | 0.450              | 0.000  | 0.069  | 0.215 | 1.072 | 0.505       |
|                    |                                               | Other non-metallic minerals              | 0.688              | 0.001  | 0.121  | 0.223 | 0.886 | 0.454       |
|                    |                                               | Furniture                                | 0.591              | 0.000  | 0.031  | 0.601 | 0.970 | 0.621       |
| Modern Industry    |                                               | Basic metals                             | 0.403              | 0.001  | 0.060  | 0.350 | 1.061 | 0.422       |
|                    |                                               | Fabricated metal products                | 0.681              | 0.000  | 0.212  | 0.072 | 0.809 | 0.480       |
|                    |                                               | Electronic components                    | 0.520              | 0.000  | 0.105  | 0.016 | 1.111 | 0.650       |
|                    |                                               | Chemicals and chemical                   | 0.515              | 0.000  | 0.045  | 0.168 | 1.068 | 0.622       |
|                    |                                               | Medical, precision and optical           | 0.526              | 0.000  | 0.082  | 0.075 | 1.085 | 0.626       |
|                    |                                               | Electrical equipment                     | 0.448              | 0.000  | 0.066  | 0.176 | 1.071 | 0.570       |
|                    |                                               | Other machinery and equipment            | 0.449              | 0.000  | 0.035  | 0.557 | 1.035 | 0.517       |
|                    |                                               | Motor vehicles, trailers                 | 0.368              | 0.002  | 0.064  | 0.334 | 1.066 | 0.387       |
|                    |                                               | Other transport equipment                | 0.400              | 0.004  | 0.057  | 0.486 | 1.053 | 0.336       |
| others             |                                               | Other manufacturing                     | 0.843              | 0.000  | 0.071  | 0.081 | 0.932 | 0.835       |

Cell color: $p < 0.1$ and adjusted-$R^2 > 0.5$. Division full name: (1) Food products; (2) textiles except apparel; (3) wearing apparel, clothing accessories and fur articles; (4) wood and products of wood and cork, except furniture; (5) pulp, paper and paper products; (6) printing and reproduction of recorded media; (7) rubber and plastics products; (8) other non-metallic mineral products; (9) furniture; (10) basic metals; (11) fabricated metal products, except machinery and furniture; (12) electronic components, computers, and visual, sounding and communication equipment; (13) chemical and chemical products, except pharmaceuticals and medicinal chemicals; (14) medical, precision and optical instruments, watches and clocks; (15) electrical equipment; (16) other machinery and equipment; (17) motor vehicles, trailers and semitrailers; (18) other transport equipment; and (19) other manufacturing (jewellery and related articles, musical instruments, sports, athletic goods, etc.).
It is assumed that this uneven rearrangement causes deviation from the typical Zipf linearity as it approaches 2017. Interestingly, traditional and modern industries are distinguished based on an exponential value of 0.6. The exponent of division 20 is similar to that of traditional industries.

The second factor leading to the rearrangement of industrial sectors is physical infrastructure. As discussion, Zipf [17] showed that a power law with a slope of one is the form is always a straight, right angle. This regularity of the Zipf distribution suggests that there is a common economy in Experiments 1 through 45,51,16,39,40. Busan, Ulsan and South Gyeongsang regions that economies act as groups that interact. However, this factor does not explain why the peaks of division 19 disappeared in 2006 is distributed in the parameter space. Looking at the distribution of Marshall $\beta_j$, the larger the exponent value, the smaller the economies of scale. Interestingly, traditional and modern industries are distinguished based on an exponential value of 0.6. The exponent of division 19 is similar to that of traditional industries.

Figure 4 shows how the combination of $\alpha_m$ and $\beta_m$ of significant divisions in 2006 is distributed in the parameter space. Looking at the distribution of Marshall $\beta_j$, the larger the exponent value, the smaller the economies of scale. Interestingly, traditional and modern industries are distinguished based on an exponential value of 0.6. The exponent of division 19 is similar to that of traditional industries.

Figure 4. Parameter space of $\alpha_m$ and $\beta_m$ in 2006.

This is presumably because division 19 is not a single industry but a division that includes traditional industries, such as precious metal processing, musical instrument
manufacturing and other categories of modern industry. Therefore, the interaction among individuals is weak, similar to that of traditional industries [54]. In the case of the Marshall coefficient \( \alpha_m \), as the values increase, the socioeconomic interaction also increases. As with the threshold of the exponent distribution, the coefficients are divided into traditional and modern industries based on a value of 1.0.

6. Discussion

6.1. Zipf Scaling

As discussed, Zipf [39] argued that a power law with a slope of -1 showed a phenomenon in which individuals behaved not as individual entities but as interacting groups, and that the slope of the power law could change; however, the form is always a straight line [38]. The urban system linearity of this study is similar to the Zipf distribution, in that it is a straight, right-down linear shape. This implies that the 22 cities that make up the urban system act as an interactive group. Particularly, the urban system alignment in large cities (such as Busan, Ulsan and Changwon) becomes a perfect straight line as it approaches 2017.

This pattern suggests that the three cities have strong interactions with the urban system. Specifically, two large peaks appear between the middle cities. This implies that the cities have their own growth engines beyond their interaction with the urban system [45, 51, 55, 56]. As the regularity of the Zipf distribution suggests that there is a common conceptual structure underlying highly complex phenomena [46], it can be inferred that Busan, Ulsan and South Gyeongsang provinces operate as mega-regions that economically interact.

Through the first experiment, we find regularity that transcends contextual constraints, such as various economic shocks [16, 39, 40]. Busan, Ulsan and South Gyeongsang provinces act as groups that interact. However, this does not explain why the peaks of medium cities occurred and what this means.

Sarkar et al. [45] suggested that, to understand the urban scaling phenomenon, it is necessary to deal with the diversity of socioeconomic behaviours, physical infrastructure and size. In line with this, two approaches are used. First, as socioeconomic behaviours are diverse, this study assumes that individual industrial sectors will be rearranged in cities through two types of agglomeration economies: Jacobs and Marshall externalities. It is assumed that this uneven rearrangement causes deviation from the typical Zipf linearity in Experiment 1.

The second factor leading to the rearrangement of industrial sectors is physical infrastructure; thus, road extension is considered in this study. The core driving force of agglomeration economies is the interaction occurring within or among cities. The force of interaction is temporally accelerated, decelerated, spatially expanded or reduced by road extension, which is a physical constraint. As an experiment in the next section, Jacobs externalities are scaled and then compared with Experiment 1.

6.2. Jacobs Scaling

In this study, the spatial level of aggregation is objectively reflected by measuring the economic activity per unit area. Consequently, as depicted in Figure 3, the city points are arranged relatively regularly around the regression line for 2006 and 2017.

When comparing Experiments 1 and 2, two conclusions may be drawn. First, the urban system of the Zipf model and the trend of the Jacobs model are both linear. If the source of the force acting as an interactive group interacting are Jacobs externalities, then the linearity of the Zipf model and the Jacobs model can be explained. Second, it is assumed that the small upward movement of the urban system annually is due to the externalities driven by the relationship between economies of scale \( \beta_j \) and socioeconomic interaction \( \alpha_j \) in the Jacobs model.

Experiment 2 explains the overall behaviour of the urban system using Jacobs externalities; however, there is a limit to the explanation of local behaviour. The linearity of the
Zipf urban system shows two peaks in the medium city group and a small hill in the small city group over time. The experiment in the next section tests the MAR model. The results are compared with those of Experiments 1 and 2.

6.3. Marshall–Arrow–Romer Scaling

The relationship between the Marshall exponents and coefficients can be summarised in three ways. First, the Marshall exponents and coefficients are inversely related. This implies that the mirror image emphasised by West [46] and the resources saved by scale economies \( \beta_m \) become the driving force of socioeconomic interaction \( \alpha_m \) [46]. Second, it implies Zipf’s [39] principle of the least effort. Inducing a high level of socioeconomic interaction with a small infrastructure investment, such as in a modern industry, implies that individual firms in the relevant industry division prefer a high level of aggregation.

If the aggregation becomes too high, diseconomies occur, such as through excessive competition. If the accumulation is too low, it is difficult to hire workers, increase transaction costs and receive benefits from the diffusion of innovation, such as acquiring new technologies. Individual firms in traditional and modern industries are constantly moving to regions where they can achieve maximum effect with minimum effort, and the Marshall exponents and coefficients show the search behaviour in the parameter space.

Third, two thresholds are identified, which are unique in the literature of spatial economic activities but have been found in the other scaling fields [57,58]. In this study, the threshold for the Marshall exponent is 0.6. If it is less than 0.6, it is a modern industry, and if it is greater than 0.6, it is a traditional industry. The threshold for the Marshall coefficient is 1.0. If it is less than 1.0, it is a traditional industry, and if it is greater than 1.0, it is a modern industry.

Sarkar et al. [45] argued that the urbanisation economy shows constant returns to scale, whereas the localisation economy shows increasing returns. In this study, the Jacobs coefficients were 1.0696 in 2006 and 1.0483 in 2017, indicating a constant value close to 1. The Marshall coefficients of the electronic components in division 12 showed values of 1.1244, 1.1110 and 1.1108 by year, suggesting that the interactions among individuals increased returns.

The limitations of Experiment 3 can be summarised in two ways. First, although the proximity within an individual city is reflected through density-based equations, there is a lack of consideration of the proximity among cities [59–62]. Second, this does not consider the size of the various spaces. If the automobile-related industries in sector 17 of the modern industry are measured in small spatial units below the city boundary, it may present interesting scaling exponents and coefficients.

7. Summary

This study comprises two parts: model development and experiments. First, four important models are developed: the function–shape linkage frame model, socioeconomic interaction coefficient, Jacobs model and Marshall model. Second, in the experiments, three hypotheses are tested for the mega-region in the south-eastern region of Korea. These consist of scaling the Zipf, Jacobs and Marshall models.

The two key academic findings from the model development and testing are as follows: First, the causes of straight lines, peaks and hills in the Zipf distribution are presented using the Jacobs and Marshall models. Second, the threshold of the level of socioeconomic interaction that can distinguish between the traditional and modern industries is determined using the Marshall coefficient. Additionally, this study can present an important theoretical framework for policymakers, in that it provides a link between spatial planning and economic policy.

Since the third industrial revolution, the proximity between firms has become increasingly important across all industries. As in the case of the industrial cluster policy, the agglomeration economy is positioned as a core principle of economic policy. However, it is difficult to find a tool that can precisely control the relationship between physical proximity
and economic proximity. The model of this study quantified the economy of scale caused by spatial planning of urban infrastructure and the socioeconomic interaction caused by the economic policy of industrial clusters, to present a frame where the two trade-off with each other. The frame structure has been simplified, so that anyone can use it easily.

A city is a type of organic organization, rather than an aggregate composed of a physical dry environment and humans. In public policy, the former is static planning and the latter is dynamic planning. This study presents a frame in which structure and change as well as form and function can be organically connected with each other. It can serve as a basis for developing theoretical and practical tools that go beyond static planning to create dynamic urban planning.

In this study, we focus on scaling as a methodology to deal with dynamics, that is, change. In book Scale, West [46] poses the challenging question: “Can we develop a conceptual framework to understand the science of cities and businesses—that is, their dynamics, growth, and evolution—in a predictable and quantitative way?” He uses various examples to explain how a system responds to a change in size. He noted that the scaling argument helps reveal a universal principle or structure and a representative one is the tipping point and phase transition. This study confirms West’s [46] scaling argument.

Despite these interesting findings and implications, further improvements are required. This study limits the scope of observation to scaling, which transcends contextual constraints. Therefore, it is possible to develop a frame model that connects the spatial shape and industrial function; however, it is necessary to consider a more precise spatial size with more micro-level economic activities.

If the theories based on scaling can be extended to location theory and the bottom-up approach, we can view not only urban growth and decline but also lifecycles, such as emergence, growth, decline and death, or the occurrence of extreme events, including the emergence of megacities or mega-regions. Finally, it is necessary to consider the tertiary industry sector, such as business services or distribution. Understanding the urban structure, dynamics and growth in modern society requires an understanding of the service industry as well as manufacturing activities.

Author Contributions: Conceptualization, J.-H.Y. and K.-W.N.; methodology, J.-H.Y.; software, J.-H.Y.; validation, J.-H.Y. and K.-W.N.; formal analysis, J.-H.Y.; investigation, J.-H.Y.; data curation, J.-H.Y.; writing—original draft preparation, J.-H.Y.; writing—review and editing, J.-H.Y.; visualization, J.-H.Y.; supervision, K.-W.N.; project administration, J.-H.Y.; funding acquisition, K.-W.N. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the [Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education] under Grant [2021R1I1A3056691]; [Kyungsung University] Research Grants in [2020-036].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analysed in this study. This data can be found here: The Microdata Service System of the Korea National Statistical Office [http://mdis.kostat.go.kr] accessed on 1 March 2021; Korean Statistical Information Service [http://www.kosis.kr] accessed on 1 March 2021.

Acknowledgments: The authors would like to thank the anonymous experts and reviewers for their valuable comments and suggestions that have improved this article.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Hwang, U.; Woo, M. Analysis of inter-relationships between urban decline and urban sprawl in city-regions of South Korea. Sustainability 2020, 12, 1656. [CrossRef]
2. Ortiz-Moya, F.; Moreno, N. The incredible shrinking Japan. City 2016, 20, 880–903. [CrossRef]
3. Yi, Z. Options for fertility policy transition in China. Popul. Dev. Rev. 2007, 33, 215–246. [CrossRef]
39. Zipf, G.K. *Human Behavior and the Principle of Least Effort*; Addison-Wesley: Cambridge, MA, USA, 1949.
40. Jiang, B.; Jia, T. Zipf’s law for all the natural cities in the United States: A geospatial perspective. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 1269–1281. [CrossRef]
41. Batty, M. *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*; The MIT Press: Cambridge, MA, USA, 2007.
42. Rosen, K.T.; Resnick, M. The size distribution of cities: An examination of the Pareto law and primacy. *J. Urban Econ.* **1980**, *8*, 165–186. [CrossRef]
43. Krugman, P. *The Self Organizing Economy*; Blackwell Publishers: Cambridge, MA, USA, 1996.
44. Kühnert, C.; Helbing, D.; West, G.B. Scaling laws in urban supply networks. *Phys. A* **2006**, *363*, 96–103. [CrossRef]
45. Sarkar, S.; Arcaute, E.; Hatna, E.; Alizadeh, T.; Searle, G.; Batty, M. Evidence for localization and urbanization economies in urban scaling. *R. Soc. Open Sci.* **2020**, *7*, 191638. [CrossRef]
46. West, G. *Scale: The Universal Laws of Life, Growth, and Death in Organisms, Cities, and Companies*; Penguin Press: New York, NY, USA, 2018.
47. Staber, U. Spatial proximity and firm survival in a declining industrial district: The case of knitwear firms in Baden-Württemberg. *Reg. Stud.* **2001**, *35*, 329–341. [CrossRef]
48. Feldman, M.P. The new economics of innovation, spillovers and agglomeration: A review of empirical studies. *Econ. Innov. New Technol.* **1999**, *8*, 5–25. [CrossRef]
49. Frenken, K.; van Oort, F.G.; Verburg, T.; Boschma, R.A. Variety and regional economic growth in the Netherlands. In *Research Series. Final Report to the Ministry of Economic Affairs*; Ministry of Economic Affairs: The Hague, The Netherlands, 2004.
50. Ning, L.; Wang, F.; Li, J. Urban innovation, regional externalities of foreign direct investment and industrial agglomeration: Evidence from Chinese cities. *Res. Policy* **2016**, *45*, 830–843. [CrossRef]
51. Rosenthal, S.S.; Strange, W.C. The determinants of agglomeration. *J. Urban Econ.* **2001**, *50*, 191–229. [CrossRef]
52. Park, B. The territorial politics of regulation under state capitalism: Regional parties and the politics of local economic development in South Korea. *Space Pol.* **2005**, *9*, 237–259. [CrossRef]
53. Park, B. Uneven development, inter-scalar tensions, and the politics of decentralization in South Korea. *Int. J. Urban Reg. Res.* **2008**, *32*, 40–59. [CrossRef]
54. Duffie, N.A. Heterarchical control of highly distributed manufacturing systems. *Int. J. Comput. Integr. Manuf.* **1996**, *9*, 270–281. [CrossRef]
55. Boonstra, B.; Boelens, L. Self-organization in urban development: Towards a new perspective on spatial planning. *Urban Res. Pract.* **2011**, *4*, 99–122. [CrossRef]
56. Fujita, M.; Krugman, P.; Mori, T. On the evolution of hierarchical urban systems. *Eur. Econ. Rev.* **1999**, *43*, 209–251. [CrossRef]
57. Hanley, Q.S.; Lewis, D.; Ribeiro, H.V. Correction: Rural to urban population density scaling of crime and property Transactions in English and Welsh Parliamentary Constituencies. Transactions of the in English and Welsh Parliamentary Constituencies. *PLoS ONE* **2016**, *11*, e0167605. [CrossRef]
58. Sotomayor-Gómez, B.; Samaniego, H. City limits in the age of smartphones and urban scaling. *Comput. Environ. Urban Syst.* **2020**, *79*, 101423. [CrossRef]
59. Devereux, M.P.; Griffith, R.; Simpson, H. The geographic distribution of production activity in the UK. *Reg. Sci. Urban Econ.* **2004**, *34*, 533–564. [CrossRef]
60. Ellison, G.; Glaeser, E.L.; Kerr, W.R. What causes industry agglomeration? Evidence from Coagglomeration Patterns. *Am. Econ. Rev.* **2010**, *100*, 1195–1213. [CrossRef]
61. Mukkala, K. Agglomeration economies in the Finnish manufacturing sector. *Appl. Econ.* **2004**, *36*, 2419–2427. [CrossRef]
62. Wennberg, K.; Lindqvist, G. The effect of clusters on the survival and performance of new firms. *Small Bus. Econ.* **2010**, *34*, 221–241. [CrossRef]