Some Empirical Regularities of Spatial Economies: A Relationship between Industrial Location and City Size

Tomoya Mori*, Koji Nishikimi†and Tony E. Smith‡

September 27, 2002

Abstract

The spatial distribution of industries and population is quite lumpy, and this lumpiness varies across industries. Nevertheless, we show using Japanese data for metropolitan areas that the locations of both industries and population are linked by surprisingly simple and persistent patterns. In addition, we show mathematically that these patterns are in turn closely related to the well known Rank-Size Rule, as applied to metropolitan areas.

Discussion paper No.551
Institute of Economic Research, Kyoto University

Keywords: Agglomeration; Hierarchy principle; Rank-Size Rule; Zipf's Law
JEL Classification: R11, R12

Acknowledgment: We thank Duncan Black, Pierre-Philippe Combes, Gilles Duranton, Masa Fujita, Robert Helseley, Seiro Ito, Hideo Konishi, Maria P. Makabenta, Se-il Mun, Henry Overman, Komei Sasaki, Akihisa Shibata, Chiharu Tamamura, Tatsufumi Yamagata, Atsushi Yoshida, and seminar participants at Boston College, Institute of Developing Economies, Kyoto University, Nagoya City University, Tohoku University and University of Tokyo for their constructive comments on an earlier version of this paper. We also thank Yoshi Kanemoto for helping us with metropolitan area data for Japan, and Noriko Tabata for helping us with county data.. This research has been supported by The Murata Science Foundation, and The Grant in Aid for Research (Nos. 09CE2002, 09730009, 13851002) of Ministry of Education, Science and Culture in Japan.

*Institute of Economic Research, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501 Japan. Email: mori@kier.kyoto-u.ac.jp. Phone: +81-75-753-7121. Fax: +81-75-753-7198.

†Institute of Developing Economies, JETRO, 3-2-2 Wakaba, Mihama-ku, Chiba, 261-8454 Japan. Email: nisikimi@ide.go.jp. Phone: +81-43-299-9672. Fax: +81-43-299-9763.

‡Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA. Email: tesmith@ssc.upenn.edu. Phone: +1-215-898-9647.

1 Introduction

The spatial distributions of both industry and population are well known to be lumpy. Most economic activities are concentrated within a fairly small portion of geographic space. In 1996, more than 80% of the workforce in Japan was contained in metro areas occupying only 30% of the available land. Even among these metro areas, the distribution of industries and population is far from uniform, and in addition, the lumpiness of industries varies across sectors. Of the 149 three-digit manufacturing industries, 45 [resp., 22] have positive employment in less than 50% [resp., 25%] of the metro areas, while 42 are ubiquitous, having positive employment in more than 90% of the metro areas.[1] Moreover, certain metro areas are seen to attract a disproportionately large number of industries, leading to great variation in industrial diversity among metro areas. The most diverse is Tokyo, with positive employment in 143 industries out of the 149, while the least diverse is Kitamizawa, with only 54 industries (and an average diversity of 97 industries across metro areas). In addition, more diverse metro area tend to have a larger populations (the rank correlation between size and diversity of metro areas is greater than 0.9).[2] The population distribution among the 118 metro areas in Japan is also very skewed. Of the total population in the metro areas, more than 60% is concentrated in the largest ten, and more than 30% in Tokyo alone.

Are there any underlying regularities governing these patterns of industrial and population locations, or could they simply happen by chance? We show using Japanese data for 1980/81 and 1995/96 that such a regularity does indeed exist. In particular, there is a strong negative log-linear relation between the number of metro areas occupied by an industry and the average size of these metro areas (shown in Figure 1 below). It is remarkable that while industries have tended over time to trickle down from bigger metro areas to (a larger number of) smaller ones during this period, this regularity has itself remained very stable. Moreover, the industrial location patterns in these periods are shown to be highly consistent with Christaller [5]'s well-known *Hierarchy Principle* of industrial location behavior, which asserts that an industry found in any given metro area will also be found in all larger metro areas.[3] These findings suggest the presence of a strong empirical regularity for industries that parallels the well-known *Rank-Size Rule* for city size distributions which asserts that if cities *in a given country* are ranked by (population) size, then the rank and size of cities are log-linearly related with the slope usually close to one.[4] The above log-linear relationship found in industrial location pattern, which we here designate

[1] For futher details on industrial classification and the definition of metro areas, see Section 2.

[2] Here, the diversity means the number of industries with positive employment.

[3] A similar but less strict hierarchical principle was suggested by Lösch [25]. So far, only informal case studies for the principle have been presented. See, e.g., Berry [2], Christaller [5], Dicken and Lloyd [6], Isard [22], Lösch [25], and Marshall [26].

[4] The Rank-Size Rule has been noted as early as Auerbach [1] and Zipf [37]. Evidence found in various countries is summarized in Rosen and Resnick [34] and Parr [32]. See also Carroll [4] for a survey of the empirical literature, and

as the *Number-Average Size (NAS) Rule*, appears in fact to be closely related to both the Rank-Size Rule for metro areas, and to Christaller [5]'s Hierarchy Principle.

In this context, the major theoretical finding of this paper is to show that whenever industrial locations are consistent with Christaller's Hierarchy Principle, the NAS Rule for industrial locations and the Rank-Size Rule for metro areas are *essentially equivalent*. In particular, if metro areas follow an appropriate Rank-Size Rule, and industries follow Christaller's Hierarchy Principle, then the NAS Rule must be a necessary consequence. However, our results also show that this equivalence is meaningful only for Rank-Size Rules with exponents *less than one*. Japan appears to be a clear case in point. But in countries with Rank-Size Rules greater than or equal to one for example, Christaller's Hierarchy Principle precludes the existence of (even asymptotic) log-linear relations of the NAS type. Hence these theoretical results may help to shed new light on the range of possible relationships between industrial location and city size.

The remainder of this paper is organized as follows. A description of data used for the analysis is given in Section 2, and is followed by a discussion of the measurement of industrial localization and diversity in Section 3 . In Section 4, we present the regularity in the relationship between the size and number of metropolitan areas in which each industry is located. In Section 5, we conduct a formal test for the Hierarchy Principle. In Section 6, the link between the regularity in the industry location and the Rank-Size Rule is investigated. We conclude with a discussion on policy and theoretical implications of our findings.

2 Population and industrial data

In this paper, we look at the location patterns of industries and of population in 1980/81 and 1995/96 (where population data is available for 1980 and 1995, while industrial data for 1981 and 1996). The individual data sources are given below.

Location and population

The geographic unit we consider is the metro area. Here we use the *Metropolitan Employment Area* (MEA) definition of metro areas developed by Kanemoto and Tokuoka [23].[5] An MEA consists of a (densely inhabited) business district that has a sufficiently large in-flow of commuters, together with the suburban counties in which these commuters reside. The business districts in an MEA consist of the central business district (CBD) with a positive net inflow of commuters, and those subcenters with significant commuter flows both to the CBD and from their surrounding counties. We identified 105

Duranton [9] for recent empirical and theoretical developments in the study of city size distributions.

[5]The MEA is comparable to the Core Based Statistical Area (CBSA) of the US. See Office of Management and Budget [29] for the definition of CBSA.

MEAs in 1980 and 118 in 1995. Both the definition of counties and the county population data employed are based on the Population Census of Japan in 1980 and 1995.[6] To make the data comparable between these two years, counties in both years are converted to those in 1995.[7]

Industries

The employment data used for the analyses in this paper are classified according to the three-digit Japanese Standard Industry Classification (JSIC) taken from the Establishment and Enterprise Census of Japan in 1981 and 1996, and applied to the respective population data in 1980 and 1995. Since industrial classifications have been disaggregated for most sectors during this 15-year period, we have attempted to reconcile the two classifications by aggregating the 1996 classification. Among the three-digit JSIC-industries, we focus here on manufacturing, services, wholesale and retail, which together include 264 industries.[8]

3 Measurement of industrial localization and diversity

To establish the regularities discussed in the introduction, it is essential to identify for each industry the set of MEAs in which that industry operates, here designated as the *industry-choice MEAs* for that industry. To determine the presence of an industry, an appropriate threshold level of employment needs to be determined. However, it is shown in Appendix 8.1 that our empirical findings are not very sensitive to the choice of threshold level. For this reason, we choose to include all MEAs with positive employment in an industry as industry-choice areas (i.e., we choose a zero threshold level). In the analysis to follow, we have chosen to designate as industry-choice areas all those areas with a positive employment level for the industry. The number of industry-choice MEAs for a given industry can then be considered as reflecting the *degree of localization* for that industry. Conversely, the number of industries found in a given MEA can be taken to reflect the *industrial diversity* of that area.[9]

There is an obvious shortcoming of this approach – namely that it ignores all information about the *size* of the industrial presence in a given MEA. For instance, a MEA with one small plant for a given industry is treated the same as an area containing many large plants in that industry. Nevertheless, we

[6] The county here is equivalent to the *shi-ku-cho-son* in the Japanese Census.

[7] Among 3375 (3370) counties, MEAs include 1329 (1564) counties in 1980 (1995).

[8] In 1981, there were a total of 319 categories (152 manufacturing, 108 services, and 59 wholesale/retail). From this set, those sectors classified as public sectors in either year, or not fittng any of the specific categories above, have been excluded. For those few sectors that are more aggregated in 1996, we follow the 1996-classification. This resulted in a final set of 125 manufacturing, 90 services, and 49 wholesale and retails industries.

[9] It should be noted that these measures differ from the more traditional measures of industrial localization (e.g., Duranton and Puga [11]; Ellison and Glaeser [13]; Krugman [24, Ch.2]) and industrial diversity in metro areas (e.g., Duranton and Puga [11]; Henderson, Kuncoro and Turner [20]). A more detailed motivation of the present measures, along with a comparison to more traditional measures, will be developed in a subsequenct paper (Mori and Nishikimi [27]).

believe that even a minimal presence of an industry in a MEA identifies that area as a *viable location* for that industry. Indeed, the literature suggests that there is a persistence of industrial locations: once an industry has successfully located in an area, the size of this industrial presence will tend to grow over time (e.g., Henderson, Kuncoro and Turner [20])[10] even in the presence of high turnover rates of establishments (e.g., Dumais, Ellison and Glaeser [8]).[11] This is also supported by our own data, which indicates that between 1981 and 1996, the number of industries in operation in more than 90% of MEAs has increased, and moreover, that there is no MEA which has completely lost the employment of more than five industries. Thus, if an industry is in operation in a MEA at a given point of time, then it is quite likely stay there in the future.

4 An empirical regularity of industrial location

As discussed above, we measure the degree of localization of a given industry by the number of industry-choice MEAs in which the employment for the industry is positive. Figure 1 shows the relationship between the number and average (population) size of industry-choice MEAs for each three-digit industry in 1980/81 and 1995/96.

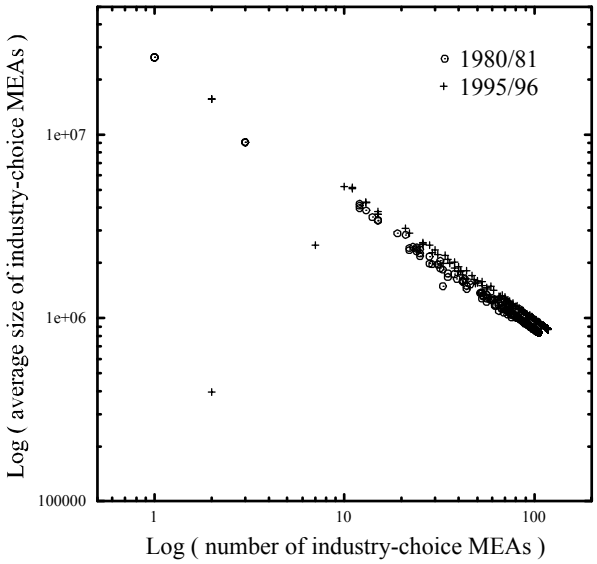

Figure 1: Size and number of industry-choice MEAs

Except for two outliers in 1995/96,[12] the relationship is clear: average size ($\overline{\text{SIZE}}$) is strongly log-linearly

[10] Henderson et. al. [20] argue that one-standard deviation increase in the proportion of 1970 local employment in a specific industry results in 30% increase in 1987 employment controling for urban size, current labor maket conditions, etc. See also Henderson [19] for a related discussion.

[11] Dumais et. al [8] report for the case of the US that nearly three-forths of plants existing in 1972 were closed by 1992, and more than a half of all manufacturing employees in 1992 did not exist in 1972.

[12] The two outliers in 1995/96 are arms-related industries: small arms manufacturing (197) and small bullet manufacturing

related to the number of industry-choice MEAs (#MEA). Ordinary least squares estimation (OLS) gives the following result,

$$1980/81 : \log(\overline{\text{SIZE}}) = \underset{(0.005164)}{7.35954}^{**} - \underset{(0.002748)}{0.7117}^{**} \log(\#\text{MEA}), \quad R^2 = 0.9961, \tag{1}$$

$$1995/96 : \log(\overline{\text{SIZE}}) = \underset{(0.003721)}{7.4170}^{**} - \underset{(0.001916)}{0.7137}^{**} \log(\#\text{MEA}), \quad R^2 = 0.9981, \tag{2}$$

where the values in the parentheses are standard errors, and ** indicates the coefficient to be significant at 1% level. In fact it should be clear by inspection that the actual significance levels are off the chart (with P-values virtually zero).

Moreover, this relationship is also seen to remain quite stable over time. By pooling the data for both periods and using time dummies, one can apply standard F-tests to evaluate coefficient shifts. The results of this analysis show that both the intercepts and slopes are significantly different (at 1% level) between the two periods. But as can be seen, the difference in slopes is very small (about 0.3% of the slope estimated from the samples in either year).

However the intercept change does reflect a significant effect, namely a differential shift between the numbers of industry-choice MEAs and the average size of those MEAs. On the one hand, there has been a dispersion of industries, as reflected by a 16% average increase in the numbers of industry-choice MEAs. On the other hand, the average size of industry-choice MEAs has increased by only 3%. Given the fact that average population size and the number of MEAs increased by 13% and 12%, respectively, one would expect to see a larger increase in the average size of industry-choice MEAs. Evidently there has been a *trickling down* of industries from bigger MEAs to a larger number of smaller MEAs, tending to lessen the effect of population growth.[13]

4.1 The NAS Rule

It should also be clear from the similarity of the slopes of these two log-linear regressions that this trickling down effect is highly structured. While both MEA sizes and numbers of industry-choice MEAs have increased, they have done so in a manner which leaves their elasticity *invariant*. In other words, the

(199). These outliers may be attributable to the fact while arms-related industries are technically in the private sector, location decisions in these industries are heavily influenced by government policies in Japan.

In addition, it should be mentioned that tobacco manufacturing has been excluded from the analysis, because it was classified in the public sector in 1981. By 1996 tobacco was privatized. Hence it is of interest to note that this industry is also an outlier in 1996, and that this again may be attributable to government policies before privatization.

[13]The diversification of smaller MEAs seems to imply convergence of industrial structure among all MEAs. However, the larger MEAs may also diversifying, keeping the relative diversity among MEAs. Namely, under the fixed industry classification, it is not possible to capture the formation of new industries which is likely to occure in the large MEAs. Notable examples are computer industries in the 80s and information technology (IT) such as internet-related industries in the 90s. In the year 2000, among the software, information processing and internet related (i.e., IT) industries found in the Yellow Page, 46.7%, 10.7% and 4.5% are located in the largest three MEAs, Tokyo, Osaka and Nagoya, respectively. Data source: Ministry of Land, Infrastracture and Transport of Japan (http://www.mlit.go.jp/). These IT industries are classfied mostly as information services (387) which includes wide variety of low-tech traditonal data processing services. Consequently, they appear as ubiquitous industry found in all MEAs.

percent change in average industry-choice MEA sizes relative to percent change in numbers of industry-choice MEAs has remained essentially constant (with a 1% increase in the number of MEAs chosen by an industry corresponding approximately to an expected decrease of .7% in the average sizes of these MEAs). Such an invariance could in principle be accounted for by simple proportional growth phenomena, such as a constant percent increase in the number of industry-choice MEAs across sectors, or a constant percent increase in average industry-choice MEA sizes across sectors. But a glance at Diagrams (a,b) in Figure 2 shows that this is not at all the case. Hence this empirical regularity appears to be much stronger, and suggests the presence of a fundamental invariance relation governing the location of industries. As a parallel to the well-known *Rank-Size Rule* for city size distributions, we designate this new relation as the *Number-Average Size (NAS) Rule* for industrial location patterns.

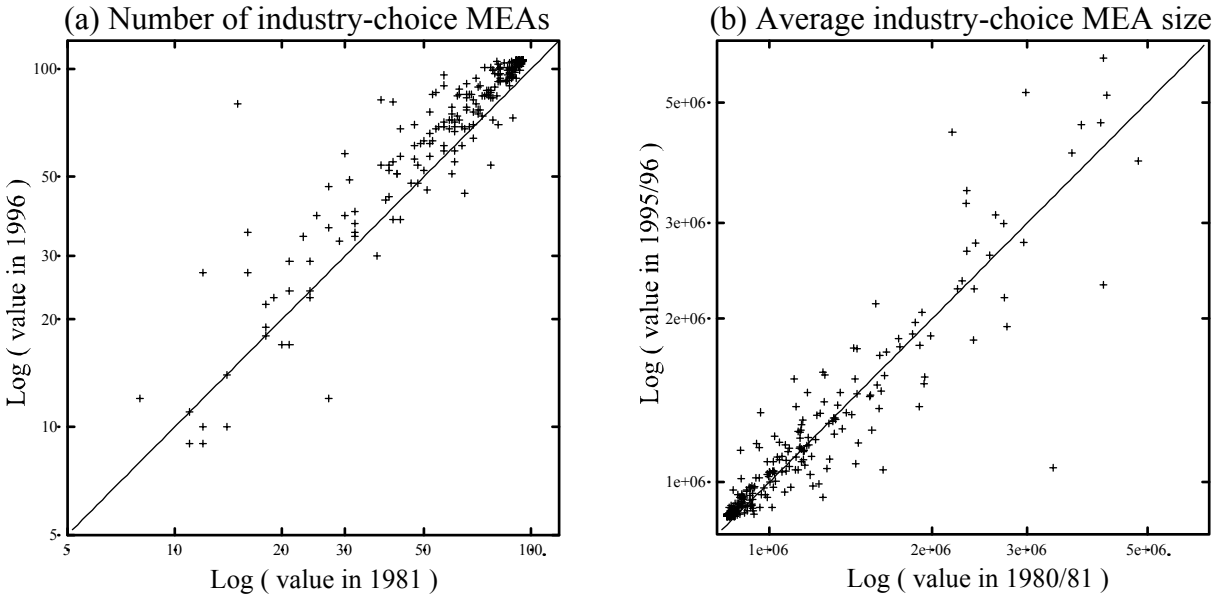

Figure 2: Change in the number and average size of industry-choice MEAs between 1980/81 and 1995/96

Two remarks are in order. The first concerns geographic aggregation and the appearance of location patterns. As we have just seen, the NAS Rule is readily apparent if MEAs are chosen to be geographic unit of aggregation. This seems reasonable in view of the fact that MEAs are *economic regions*, essentially embodying the urban economic notion of "cities" (i.e., commuting areas in which individual firms and consumers share a common urban environment). Thus, if population size is a significant determinant of industrial location, then it is natural to expect that MEAs should constitute the appropriate geographic unit.[14] Not surprisingly, if the geographic unit is chosen arbitrarily, then the NAS Rule becomes am-

[14]Duranton and Overman [10] provide evidence for the UK manufacturing that the geographic extent of establishment concentration roughly corresponds to that of an metro area.

biguous. Diagrams (a,b) in Figure 3 show the 1995/96 case, where the geographic unit is chosen to be "county"and "prefecture", respectively.[15] We can see that the clear pattern of Figure 1 now disappears. This suggests that analyses of industrial location based on inappropriate choices of geographic units may produce misleading results.

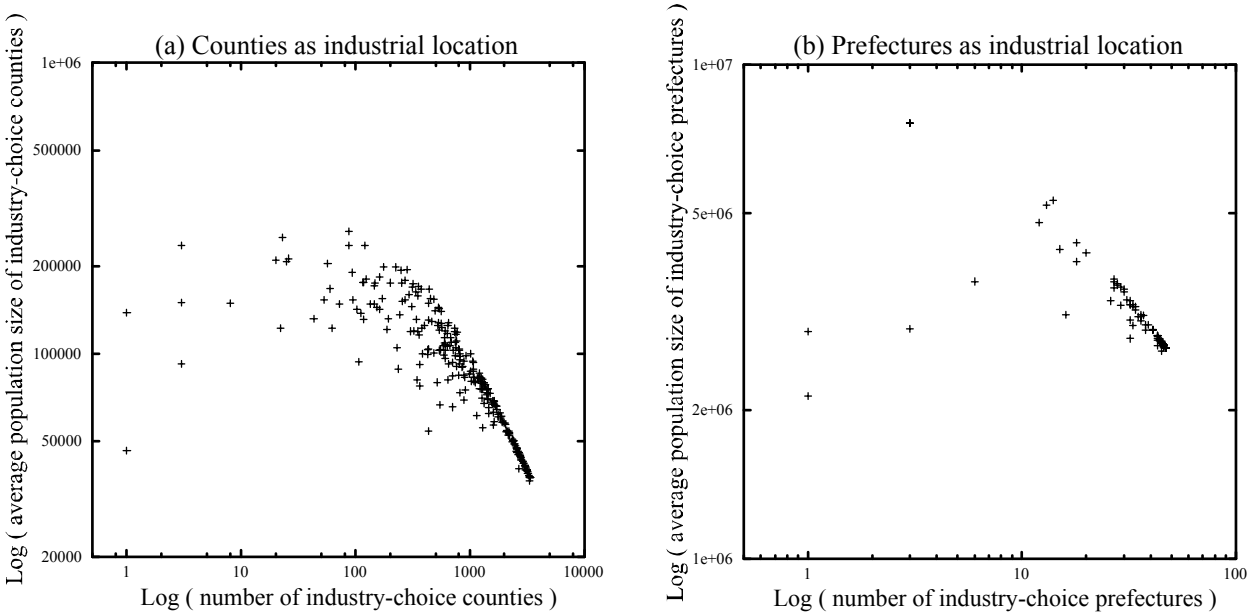

Figure 3: Alternative geographic units

The second remark concerns our choice of threshold employment in an MEA (i.e., a zero-employment threshold level) to indicate the presence of an industry. To show that the NAS Rule for Japan is robust with respect to the choice of threshold level for defining "presence," we have replotted the data for 1995/96 in Appendix 8.1 using a range of alternative threshold levels. These plots show that the NAS Rule is in fact quite robust with respect to the choice of a threshold level.

4.2 A Test of the NAS Rule

To further clarify the significance of this rule, it is of interest to consider what the relationship between numbers and average sizes of MEAs for each industry would look like if industrial locations were in fact random. To do so, we begin by observing that any process of randomly relocating industries necessarily leads to some ambiguity concerning the resulting "sizes" of MEAs. For example, it makes little sense to relocate a steel plant without considering the employment needed to operate that plant. Hence we

[15]In the figure, as in the case of MEA, the county [resp., prefecture] in which a given industry is present is called industry-choice county [resp., prefecture] of the industry.

consider it more appropriate to relocate employment along with industries. More specifically, we here treat the employment of each industry in an MEA as a single *employment cluster* for that industry, and study the effects of randomly reallocating these clusters among MEAs.[16] This relocation process amounts to randomly "regrowing" the industrial employment structure of each MEA.[17] Ideally, one would also like to relocate the entire population tied to a given industrial cluster (such as the family members of industrial workers). But rather than make arbitrary assumptions as to the sizes of these populations, we choose to focus only on the employment sizes of industrial clusters. Hence in any random reallocation of industrial clusters, we take the resulting *size* of each MEA to be simply the total size of all employment clusters assigned to that MEA.

In this context, we now consider the relationship between the number of industry-choice MEAs and the *average employment size* of those MEAs for 1995/96, as shown by the scatter plot of "+" symbols in Figure 4. Notice that this plot closely resembles the same plot in Figure 1 using average population sizes of MEAs, and hence that our present restriction to employment sizes evidently has little effect on the NAS Rule itself. To test this actual pattern against random location patterns, we generated 1000 random reallocations of all employment clusters, and calculated the corresponding average employment sizes of the MEAs to which each industry was assigned. The results of these calculations have been summarized in Figure 4, by plotting the means (solid horizontal line) together with the 5 and 95 percentile points (broken lines) for the sampling distributions of average MEA sizes for each industry (ordered by their numbers of industry-choice MEAs).[18] These results show quite convincingly that the actual pattern is highly nonrandom, and in particular that the average employment sizes of MEAs for each industry are vastly greater that would be expected by random agglomerations of industries. More formally, if the above sampling distribution is taken to represent the null hypothesis of "purely random industrial agglomerations", then our results provide strong evidence against this null hypothesis.

In fact, the observed average employment sizes not only exceed those of all 1000 samples in every case, but actually approach their *maximum possible levels*, as indicated by the "upper bound" curve in the Figure. Here, for each number of industry-choice MEAs, n, the value plotted corresponds to the actual average size of the largest n MEAs.[19]

[16] Here we sample without replacement, so that distinct employment clustered are randomly assigned to distinct MEAs.

[17] Note that the size of an MEA (both in terms of population and area) is in reality *not* independent of industrial location. Rather, it is at least in part determined by processes of economic agglomeration. For example, the dramatic growth of Tokyo has been significantly influenced by the decisions of many industries to locate there. Thus, it seems to us to be more natural *not* to take the total size of an MEA as given when industrial location is randomized. While there are of course many alternative randomization schemes, we expect the basic results presented here would be the same.

[18] The results do not basically change by increasing the number of samples above 1000.

[19] Strictly speaking, this upper bound is relevant only for the actual location pattern (the '+' points) and not for the 1000 random samples, since the latter involve somewhat different MEA sizes.

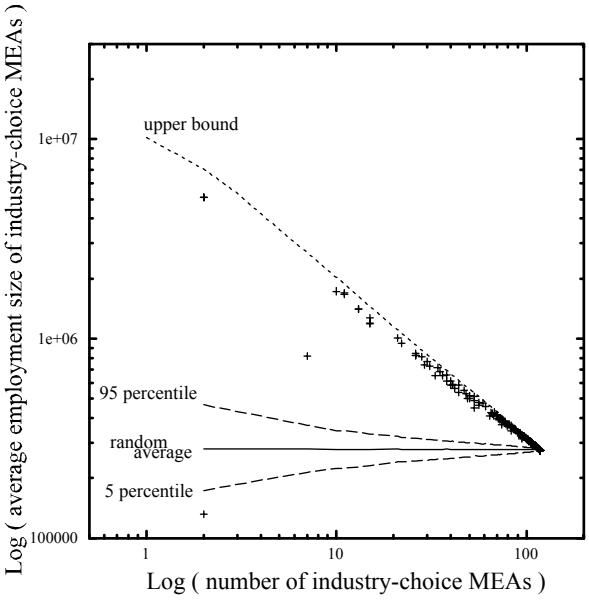

Figure 4: A Test of the NAS Rule using 1995/96 data

Hence the actual pattern of average-employment sizes is not only strongly log linear in n, but is seen to be well approximated by its upper bound – which turns out also to be strongly log linear in n. In the next section we introduce Christaller's Hierarchy Principle, which will help to account for this upper-bound approximation. In Section 6 below, we return to the log linearity property of this upper bound, which turns out to be closely related to the log linearity of the more well-known Rank-Size Rule.

5 Hierarchy Principle

Recall from the Introduction that Christaller's Hierarchy Principle asserts that industries found in any given metro area will also be found in all larger metro areas If this principle holds exactly, then it should be clear that for any given number of industry-choice MEAs, the average size will always be as large as possible. Hence the closeness of average employment sizes to the "upper bound" in Figure 4 already suggests a strong confirmation of this principle. Moreover, if average employment size is replaced by average population size, then essentially the same picture emerges (see Fig. 9). But such observations do not constitute a formal statistical test of this principle. Hence, as in the case of the NAS Rule above, it is desirable to construct an appropriate null hypothesis of "purely random industrial hierarchies" and to test this hypothesis against the observed data.

In doing so, however, one encounters several problems. First of all, the Hierarchy Principle is defined

with respect to a given population distribution among MEAs. But any attempt to relocate population along with industries will necessarily change that distribution, and hence create certain ambiguities in the interpretation of the Hierarchy Principle itself. Moreover, as was seen in the test of the NAS Rule above, there is no obvious way to relate the relocation of population to that of industry. Hence it appears that there is a need for an alternative approach to testing this principle. The second problem encountered is that unlike the NAS Rule above, where the values of average employment sizes of industrial-choice MEAs provide natural test statistics, the presence or absence of hierarchies is not measurable by any single numerical value. Hence the construction of statistical tests is somewhat more problematic. We now address these two problems in turn.

5.1 Hierarchies in terms of industrial diversity

As in the case of the NAS Rule above, one could replace populations with employment clusters, and redefine industrial hierarchies in terms of total employment size. However, a major shortcoming of this approach in our view is that it ignores regional structure. For example "Tokyo" and "Osaka" are treated simply as abstract locations when reallocating industries. Hence we believe that it is more appropriate to construct tests which preserve regional structure and attempt to identify the presence or absence of industrial hierarchies within that structure. With this in mind, we choose here to focus on the given structure of *industrial diversity* among regions rather than population per se. This approach has several advantages. First of all, as will be seen below, this leads to a very natural random sampling procedure for testing the Hierarchy Principle. In particular, if m industries are present in a given MEA then one may regard this MEA as having m "slots" to which industries can be randomly assigned. This ensures, for example, that Tokyo and Osaka will continue to be major industrial centers in each random sample.

The second advantage of this approach is that industrial diversity of MEAs is in reality very closely related to their population size. Figure 5 shows this relationship for MEAs in 1995/96.[20]

Note that while the three largest MEAs (Tokyo, Osaka and Nagoya) are almost fully diversified, this is in part due to the industrial classification system used. In particular, the three-digit JSIC does not allow us to distinguish between the industrial compositions of these MEAs (refer to footnote 5). For this reason, we have chosen to exclude them from our statistical comparison between size and diversity of MEAs. The resulting log-linear regression for the 1995/96 data is shown below.[21]

$$\log(\mathrm{DIV}) = \underset{(0.04643)}{1.516}\ ^{**} + \underset{(0.008305)}{0.1374}\ ^{**} \log(\mathrm{SIZE}), \quad R^2 = 0.7025. \tag{3}$$

[20] A similar plot can be obtained for 1980/81.

[21] There is no significant difference in the estimated coefficients between 1980/81 and 1995/96.

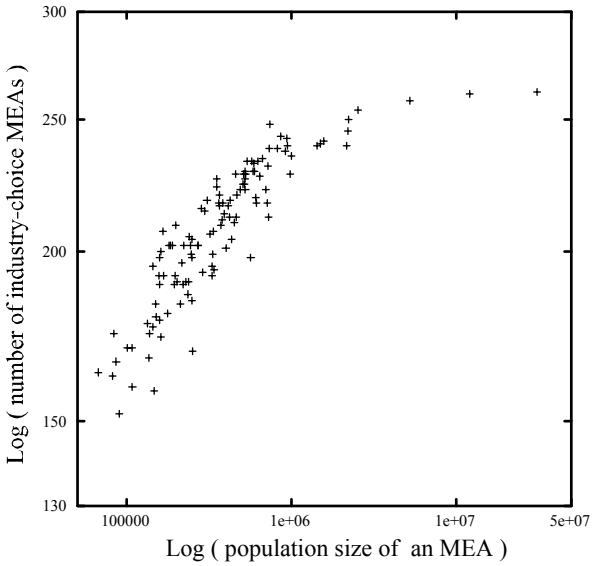

Figure 5: Population size and industrial diversity of MEAs

While the R^2 indicates that some unexplained dispersion remains, nevertheless the significance of the slope coefficient rivals that of the NAS Rule above (with P-value virtually zero). In addition, the rank correlation between the size and diversity is greater than 0.9 for both 1980/81 and 1995/96. Thus, there is close agreement between the rankings of MEAs in terms of diversity and population size.[22]

For these reasons, we now focus on industrial diversity, and hence redefine the Hierarchy Principle for testing purposes as follows. An industrial location pattern is said to satisfy the *Hierarchy Principle* iff industries found in a given MEA are also found in all MEAs with diversities at least as large.[23] Following

[22]It is to be noted that this correlation between size and diversity of MEAs appears to be stronger than similar correlations obtained using alternative indices reported in the literature. Three are worth mentioning. First, the *Herschman-Herfindahl index* (HHI) is often used (e.g., Henderson [19]; Henderson, Kuncoro and Turner [20]), as an "inverse" measure of industrial diversity of MEAs, and is defined to be the sum of the squared employment shares for each industry in a given MEA. Next, Duranton and Puga [11] measure diversity of metro areas by a *relative diversity index* (RDI) defined to be the inverse of the sum of the absolute differences between the employment share of each industry in a metro area and its corresponding national employment share. The rank correlation between size and diversity for our MEAs in 1995/96 is only 0.37 when diversity is measured by HHI, and 0.55 when the RDI is used.

However, these low correlations may be in part accounted for by the particular properties of these indices. With respect to HHI, it should be noted that this index is influenced by the difference in labor requirement across industries. For instance, given the number of industries in operation in a MEA, the MEA is evaluated as more diversified if the employment share for each industry is similar. Thus, for instance, a MEA which happens to have an industry with large labor requirement will look less diversified. Turning next to RDI, it is easy to see that according to this index the largest MEAs are not the most diverse areas since they are "more diversified" than the nation as a whole. Hence this index is clearly at odds with our empirical findings. [See Mori and Nishikimi [27] for a more detailed discussion of diversity indices based on employment shares]

[23]It should be noted that from a formal viewpoint, this principle might be designated as the *Weak Hierarchy Principle,* since it is strict weakening of the classical Hierarchy Principle. In particular, if industries in a given MEA always occupy MEAs populations at least as large, then these MEAs will always have at least as many industries, i.e., at least as much diversity. Thus the classical Hierarchy Principle implies this Weak Hierarchy Principle. However the converse is false. For example if it were true that industries in a given MEA always occupy MEAs with populations at least as *small*, then while the classical Hierarchy Principle obviously fails, the Weak Hierarchy Principle continues to hold (since MEAs with smaller population will always have at least as much diversity as those with larger population). However, for sake of simplicity, we choose to ignore this formal distinction and refer to this weaker version as the "Hierarchy Principle".

Christaller [5], we call an industry which locates in a smaller [resp. larger] number of MEAs a *higher-order* [resp., *lower-order*] industry, and call a more [resp., less] diversified MEA a *higher-order* [resp., *lower-order*] *MEA*.[24]

In these terms, the presence of an industrial hierarchy can be observed for MEAs in Japan in 1996 as in Figure 6 below.[25] Here MEAs are ordered on the horizontal axis by their industrial diversities (i.e., numbers of industries they contain), and industries are ordered on the vertical axis by their numbers of industry-choice MEAs. Hence each point "+" in the figure represents the event that the MEA in that column contains the industry in that row

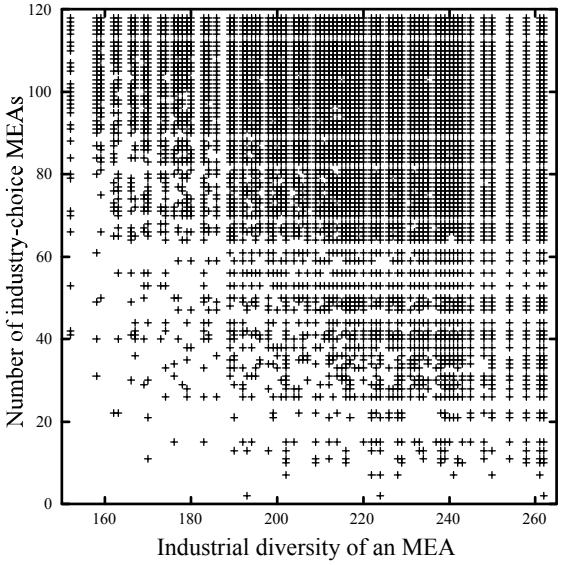

Figure 6: Hierarchy principle (1995/96)

Notice that the points are more sparse near the southwest corner, meaning that the industries with a smaller number of locations are found mainly in MEAs with large industrial diversity. On the other hand, MEAs with small industrial diversity tend to have more ubiquitous industries (i.e., those locating in a large number of MEAs).

[24]Higher-order industries are typically either subject to large scale economies (e.g., coke, briquette, blast furnace manufacturing, petroleum refining) or highly specialized (e.g., fur, leather, surveying equipment, fireproof product, spectacle manufacturing, special school education services, social and cultural science research services). Lower-order industries are subject to high tranport cost in the general sense. Examples of these are the manufacture and wholesale/retail of perishable products (e.g., meat and dairy food, vegetable and fruit food), that of heavy product (e.g., stone and related product, cement), and services/retails of frequent use (e.g., attorney services, department strores, automobile maintenance, drug and cosmetic retail). The detailed result is available from the authors upon request.

[25]A similar plot can be obtained for 1981.

5.2 A test of the Hierarchy Principle

While the observations regarding Figure 6 in the previous section are highly suggestive, they in no way constitute a *test* of the Hierarchy Principle. Moreover, it should be clear that by pure chance alone, any given industry is more likely to be in those MEAs with higher diversity, i.e., with greater numbers of industries. Hence it is important to discriminate statistically between these random hierarchical effects and genuine hierarchies. As stated above, our approach here is to construct an appropriate test statistic for measuring "degree of hierarchy," and to formulate an appropriate null hypothesis of "purely random hierarchies" in terms of this statistic.

5.2.1 The hierarchy-share statistic

We begin with the construction of a test statistic. To do so, note first that each cell containing a + in Figure 6 above can be viewed as part of a larger "hierarchy event" if all cells to the right also contain a +, i.e., if that industry is contained in all MEAs with diversities at least as large. Hence if we assign a value 1 to each cell with a + whenever it is part of a hierarchy event, and assign a value 0 otherwise, then we can use these binary outcomes to construct an appropriate test statistic. In particular, the total of these values must lie between zero and the total number of +'s. Hence dividing by this maximum total, we obtain a natural test statistic, P, representing the fraction of hierarchy events that occur. This *hierarchy share,* P, must always lie between 0 and 1, and must achieve its maximum value, $P = 1$, exactly when the Hierarchy Principle holds. Thus P constitutes a natural test statistic for measuring "closeness" to the Hierarchy Principle.

To formalize these ideas we begin with a set of industries, $i \in \mathbf{I} = \{1, \ldots, I\}$, and a set of MEAs, $r \in \mathbf{R} = \{1, .., R\}$. The industry-choice MEAs for industry i then correspond to some subset, $\mathbf{R}_i \subseteq \mathbf{R}$, and the number of these MEAs is given by the cardinality, $n_i = \#\mathbf{R}_i$. We assume (as in Figure 6) that industries are ranked by their choice numbers so that for any $i, j \in \mathbf{I}$, $i > j \Rightarrow n_i \geq n_j$. Similarly, if the number (diversity) of industries in a given MEA, r, is denoted by D_r, then we assume that MEAs are ranked by these values so that for all MEAs, $r, s \in \mathbf{R}$, $r > s \Longrightarrow D_r \leq D_s$. Hence MEAs are assumed to be ordered from the largest diversity ($r = 1$) to the smallest diversity ($r = R$), and the ordered set, $D = (D_r : r \in \mathbf{R})$, is designated as the corresponding *diversity structure.* In this context, each cell of Figure 6 corresponds to a pair (n_i, D_r), which we now represent simply by (i, r).

Next we define a *location indicator*, x_{ir}, for each industry, $i \in \mathbf{I}$, and MEA, $r \in \mathbf{R}$, by

$$x_{ir} = \begin{cases} 1 & , \text{ if industry } i \text{ locates in } r \\ 0 & , \text{ otherwise} \end{cases} . \tag{4}$$

For the given *observed diversity structure*, $D_0 = \{D_{0r} : r \in \mathbf{R}\}$, each vector, $x = (x_{ir} : i \in \mathbf{I},\ r \in \mathbf{R})$, of

indicator variables satisfying $\sum_{i\in\mathbf{I}} x_{ir} = D_{0r}$ for all $r \in \mathbf{R}$, is designated as a *feasible location pattern* for D_0. In particular, the *observed location pattern* given by the empirical data is obviously feasible, and is denoted by $x^0 = (x^0_{ir} : i \in \mathbf{I},\ r \in \mathbf{R})$.

For each feasible location pattern, x, and ir-pair, we let $S_{ir} = \{s \in \mathbf{R} : D_s \geq D_r\}$, then one can define the *hierarchy event,* $H_{ir}(x)$, by

$$H_{ir}(x) = \begin{cases} 1 & , \text{ if } x_{is} = 1 \text{ for all } s \in S_{ir} \\ 0 & , \text{ otherwise} \end{cases} , \tag{5}$$

so that the event *occurs* when $H_{ir}(x) = 1$. Observe that the event $H_{ir}(x)$ is only *feasible* (i.e., can only occur) for ir-pairs with $x_{ir} = 1$. If the number of *feasible hierarchy events* is denoted by $h = \sum_{r\in\mathbf{R}} D_{0r} = \sum_{ir} x_{ir}$, then the fraction of hierarchical events in x which occur (i.e., are consistent with the Hierarchy Principle) is given by

$$P(x) = \frac{1}{h} \sum_{ir} H_{ir}(x). \tag{6}$$

and is designated as the *hierarchy share* for x. In particular, the value, $p_0 = P(x^0)$, denotes the *observed hierarchy share* derived from the given data.

5.2.2 The Testing Procedure

In term of this statistic, we next formulate an appropriate procedure for testing the Hierarchy Principle as follows. In the absence of any additional hierarchical structure, our fundamental hypothesis is that all location patterns consistent with the observed diversity structure, D_0, should be *equally likely.* Hence if the set of feasible location patterns for D_0 is denoted by

$$\mathbf{X}_0 = \left\{ x = (x_{ir} : i \in \mathbf{I},\ r \in \mathbf{R}) : \sum_{i\in\mathbf{I}} x_{ir} = D_{0r}\ ,\ r \in \mathbf{R} \right\} \tag{7}$$

and if the elements of $\mathbf{X}_0$ are regarded as possible realizations of a random vector, $X = (X_{ir} : i \in \mathbf{I},\ r \in R)$, then our null hypothesis is simply that

$$H_0 : X \textit{ is uniformly distributed on } \mathbf{X}_0 \tag{8}$$

In theory, one would then like to derive the distribution of the hierarchy-share statistic, $P(X)$, and state this hypothesis directly in terms $P(X)$. The appropriate test of the Hierarchy Principle would then amount to determining how likely it is that the observed hierarchy share, p_0, came from this statistical population. However, in the present case it is far simpler to simulate many random draws from X under H_0, and then construct the corresponding *sampling distribution* of $P(X)$. In particular, to draw a random sample x from X under H_0, one may proceed as follows:

(i) For each MEA, $r \in \mathbf{R}$, randomly sample a subset of D_{0r} industries from $\mathbf{I}$ without replacement [say by randomly permuting the set of labels $\mathbf{I} = \{1, .., I\}$, and taking the sample to be the first D_{0r} elements from this list].

(ii) Set $x_{ir} = 1$ if industry i is in the sample drawn for r, and $x_{ir} = 0$ otherwise.

It should be clear that this procedure will automatically yield a feasible location pattern $x \in \mathbf{X}_0$, and that all such patterns will be equally likely. So to construct a set of M independent samples of the hierarchy-share statistic, say $\{P_m : m = 1, .., M\}$, under H_0 :

(iii) Construct sample location patterns, $\{x^m : m = 1, .., M\}$ using (i) and (ii), and set $P_m = P(x^m)$, $m = 1, .., M$.

Finally to carry a test of the Hierarchy Principle, observe that one may estimate the cumulative distribution function, $F(p) = \Pr(P \leq p)$, under H_0 by

$$\hat{F}(p) = \frac{1}{M} \#\{m : P_m \leq p\} \tag{9}$$

where $\#\{m : P_m \leq p\}$ represents the number of samples with $P_m \leq p$. Following standard testing procedures, one can then estimate the appropriate *P-value* for the test by

$$\widehat{\Pr}(P \geq p_0) = 1 - \hat{F}(p_0) \tag{10}$$

This estimates the probability of getting a hierarchy share as large as the observed fraction, p_0, under H_0. If M is sufficiently large to ensure a good approximation in (10), and if this P-value is sufficiently small, say $< .01$, then one can safely conclude that the actual data is much closer to the Hierarchy Principle than would be expected under H_0.

In summary, this testing procedure appears to offer three important advantages: (i) the test is based on a simple summary statistic that reflects the "degree of closeness" of location patterns to the Hierarchy Principle, (ii) it requires only simple random sampling for the required simulations, and (iii) it is completely nonparametric and makes no appeal to asymptotic distribution theory.

5.3 Test results

Within this testing framework, we calculated the hierarchy shares for 1000 randomly generated location patterns under H_0. The sampling distribution (histogram) of P obtained from this sample is shown in Figure 7 (with sample mean .148 and standard deviation .006) In addition, the observed hierarchy shares for the empirical location patterns in 1981($p_0 = 0.701$) and 1996 ($p_0 = 0.716$) are indicated by the vertical broken lines in the figure. As it is obvious from the figure, these hierarchy shares are vastly higher than

that those of the random samples. In fact, there is no hierarchy share for a random sample which comes even close to that of the empirical location patterns, and the estimated P-value is virtually zero in each case.

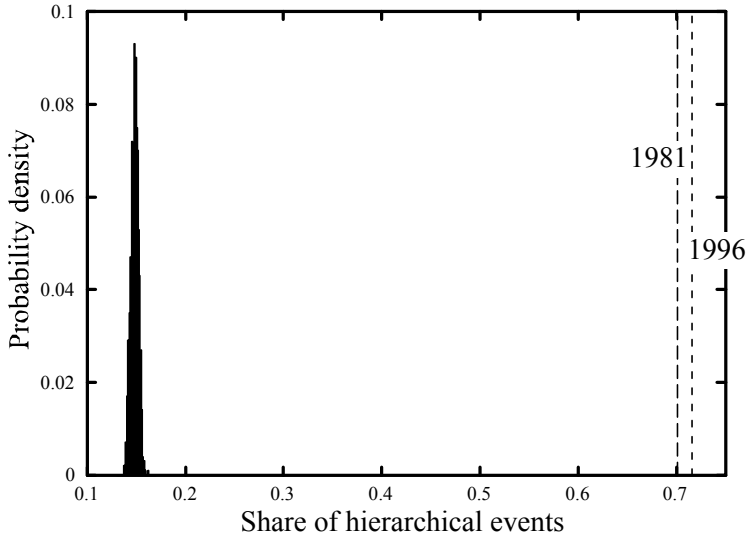

Figure 7: Sampling distribution of hierarchy shares under H_0

5.4 Alternative testing procedures

As mentioned above, there are many possible approaches to testing this Hierarchy Principle. Here we mention two other possibilities that we have explored. First observe that our emphasis in the above approach has been to preserve regional structure in terms of industrial diversity. Here industries themselves were treated as completely interchangeable under the null hypothesis, and were randomly assigned to MEAs in any manner consistent with the given levels of industrial diversity. An alternative approach is to focus on industries, and to preserve the localization structure of industries rather than the diversity structure of MEAs. In this case one would start with an *observed localization structure*, $n^0 = \{n_i^0 : i \in \mathbf{I}\}$, of industry-choice MEA numbers for each industry, and [as a parallel to (7) above] would define the set of *feasible location patterns*, $\mathbf{X}^0$, consistent with n^0 by

$$\mathbf{X}^0 = \left\{ x = (x_{ir} : i \in \mathbf{I},\ r \in \mathbf{R}) : \sum_{r \in \mathbf{R}} x_{ir} = n_i^0\ ,\ i \in \mathbf{I} \right\} \tag{11}$$

In this context, the appropriate null hypothesis, H^0, now take the form:

$$H^0 : X \textit{ is uniformly distributed on } \mathbf{X}^0 \tag{12}$$

One can then construct random samples from $\mathbf{X}^0$ in a manner paralleling the procedure above. Here for each industry, $i \in \mathbf{I}$, one randomly samples a subset of n_i MEAs and defines $x_{ir} = 1$ iff r is in the assignment for industry i. The construction of hierarchy shares, $P(X)$, is the same as before. Hence the appropriate test of H^0 amounts to estimating the P-value of the observed hierarchy share, p_0, above under the new sampling distribution obtained from these randomly generated hierarchy shares.

We have carried out such a test, and the results are shown in Figure 8. Here it is clear that observed hierarchy shares, p_0, for both 1981 and 1996 continue to be vastly greater than any of the randomly sampled shares, and hence that this test also provides strong evidence in favor of the Hierarchy Principle versus H^0. However, one interesting difference between these two test results is that the sampling

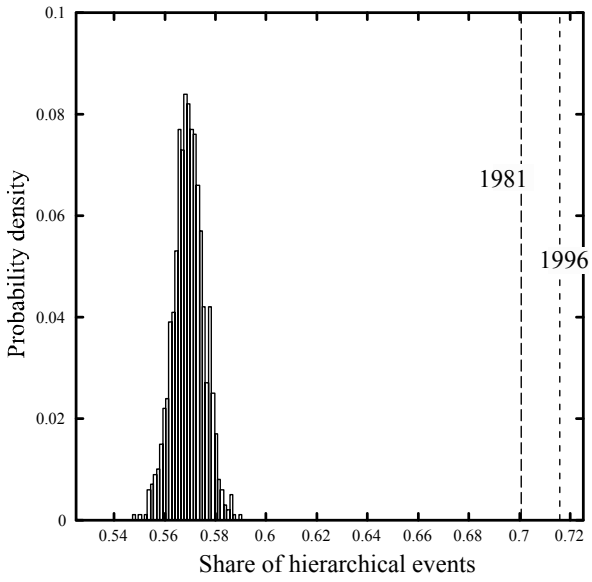

Figure 8: Sampling distribution of hierarchy shares under H^0

distribution of hierarchy shares is much larger under H^0 than under H_0. This at first seems surprising since H^0 ignores all regional structure, and treats MEAs as completely interchangeable. However, further reflection shows that the key difference here relates to ubiquitous industries. In Japan, approximately 52% of the industries are located in more than 90% of all MEAs. Hence when the observed location structure, n^0, is held fixed, there must necessarily be a large number of hierarchy events associated with these ubiquitous industries in every random sample. It can thus be inferred that when location structure is randomized (as under H_0), this leads to a much smaller range of n_i-values than is actually observed. [By looking at the average percent occupancies for industries over the 1000 samples shown in Figure 7, it turns out that the *most ubiquitous* industry occupies less than 90% of all MEAs, and the *most localized* industry occupies more than 60% of all MEAs.] Conversely, if the observed diversity structure, D_0, is

randomized (as under H^0), our results show that this leads to a much smaller range of D_r-values than is actually observed. Hence *both MEAs and industries are in reality much more heterogeneous than would be expected in these "randomized" worlds.*

A second approach to testing is motivated by the fact that the hierarchy shares defined in (5) are in fact rather stringent. In particular, consider two industries, $i, j \in \mathbf{I}$, that each are located in all but one MEA. If the MEA not containing i were one of the smallest, then this would generate many hierarchy events for industry i. But if the MEA containing j happened to be one of the largest, then there would be very few hierarchy events for j. Hence even though both industries are located in almost the same set of MEAs, one is highly consistent with the Hierarchy Principle while the other is not. This extreme example suggests that the 0-1 hierarchy-event measure, H, should perhaps be replaced with a more graduated measure of "closeness" to pure hierarchy events. One obvious measure of this type looks at the *fraction* of possible location events that actually occur "at or above" any given event. More precisely, if for any location event, x_{ir}, we let $s_{ir} = \#S_{ir}$, then the desired *fractional hierarchy event*, $\overline{H}_{ir}(x)$, is defined for each location pattern x by

$$\overline{H}_{ir}(x) = \frac{1}{s_{ir}} \sum\nolimits_{k \in S_{ir}} x_{ik}$$

One can then average these values to produce an overall test statistic, $\overline{P}(x)$, of *fractional hierarchy shares*. A test of H_0 using $\overline{P}$ rather than P produces results that are qualitatively the same as those above. As expected, the values of $\overline{P}$ are uniformly higher than those of P, but the observed values, $\overline{p}_0$, are always well above all sampled values. Hence this test provides even further support for the Hierarchy Principle.

6 Industrial location and the Rank-Size Rule

Given the strong evidence above for both the NAS Rule and the Hierarchy Principle, we now consider a possible relationship between these two empirical regularities. As stated in the introduction, it turns out that in the presence of the Hierarchical Principle, the NAS Rule is essentially identical to perhaps the most widely known empirical regularity in all of economic geography, namely the *Rank-Size Rule* (or Zipf [37]'s Law) for city size distributions. This rule asserts that if cities *in a given country* are ranked by (population) size, then the rank and size of cities are log-linearly related:

$$\log SIZE = \sigma - \theta \log RANK. \tag{13}$$

Moreover, the estimate of θ is usually close to one. The most complete cross-country comparison for the relationship is conducted by Rosen and Resnick [34] who reported that in 1970 the estimates of θ for 44

countries lie between 0.51 and 1.24 with an average value 0.90 and standard deviation, 0.14.[26] It is to be noted, however, that the definition of "city" in their analysis is the administrative city, which often fails to represent cities in economic sense. They show for countries where data for metro area exists, the estimates of θ are even closer to 1 when metro areas are used instead of administrative cities. For our Japanese MEAs, we have obtained the estimated values, $\theta = 0.95$ ($R^2 = 0.95$) for 1980/81 and $\theta = 0.99$ ($R^2 = 0.95$) for 1995/96. Although there is no definitive evidence for the Rank-Size Rule, the recent literature suggests that the observed rank-size relationships are too close to this rule to be completely dismissed as irrelevant (e.g., Black and Henderson [3], Dobkins and Ioannides [7], Ioannides and Overman [21]).

To establish a relationship between this rule and the NAS Rule in the presence of the Hierarchy Principle, it is useful to begin with a simple formulation using a continuum of population units (MEAs). Here the essence of the connection is mathematically transparent, though the notions of "population size" and "rank" are somewhat tenuous. We then develop a more realistic (but mathematically more complex) version of this relationship in terms of discrete population units.

6.1 A continuous model

Consider a continuum of *population units* (MEAs) on the interval $\mathbf{R} = [0, R]$, where each population unit, $r \in \mathbf{R}$, is ranked by its *size*, $\rho(r)$, with $r < r' \Leftrightarrow \rho(r) > \rho(r')$. The largest population unit is represented by rank 0, the smallest by rank R, and the total number (mass) of population units by $R = \int_{\mathbf{R}} dr$. Consider also a set of *industry types*, $i \in \mathbf{I}$, with each industry i occupying a (measurable) subset, $\mathbf{R}_i$, of the population units in $\mathbf{R}$, and let the number (mass) of those units be denoted by $n_i = \int_{\mathbf{R}_i} dr$. Assuming that ρ is continuous on $\mathbf{R}$, it follows that *average size*, $\overline{R}_i$, of those units occupied by i is given by

$$\overline{R}_i = \frac{1}{n_i} \int_{\mathbf{R}_i} \rho(x) dx. \tag{14}$$

In such an economy, $\mathbf{E} = (\mathbf{R}, \mathbf{I}, \rho, n)$, the linear-linear relationships in expressions (1) and (2) now take the form

$$\ln(\overline{R}_i) = a - \beta \ln(n_i), \qquad a, \beta > 0 \tag{15}$$

where for convenience we ignore any random factors and use only simple equalities. This log-linear relation, which can also be written as

$$\overline{R}_i = \alpha n_i^{-\beta}, \qquad \alpha = \exp(a) > 0, \beta > 0 \tag{16}$$

[26] In Rosen and Resnick [34], $\log RANK$ is regressed against $\log SIZE$.

is now designated (in a manner similar to Section 4.1) as the *Number-Average Size* (*NAS*) *Rule* with scale factor α and exponent β for industrial locations in economy $\mathbf{E}$.

The *Hierarchy Principle* for economy $\mathbf{E}$ then asserts that for each industry type, $i \in \mathbf{I}$, the population units occupied by i are an *interval* in $\mathbf{R}$ of the form $\mathbf{R}_i = [0, r_i]$, where r_i denotes the rank of the smallest population units occupied by i. Hence the total number, n_i, of units occupied by i is now given by r_i, so that the average size in (14) is given by:

$$\overline{R}_i = \frac{1}{r_i} \int_0^{r_i} \rho(r) dr. \tag{17}$$

Finally, we recall that population units, $r \in \mathbf{R}$, are said to satisfy a *Rank-Size Rule* with scale factor s and exponent θ if and only if :

$$\rho(r) = sr^{-\theta}, \qquad s, \theta > 0. \tag{18}$$

[where $s = \exp(\sigma)$ in expression (13) above]. With these definitions, our main result is to show that in the presence of the Hierarchy Principle, the NAS Rule for industrial locations is essentially equivalent to the Rank-Size Rule for population units, and in addition that the exponents β and θ must be the same and less than one. In particular, if we now say that $\mathbf{I}$ is a *full set of industries* for economy $\mathbf{E}$ whenever there exists for each population unit, $r \in \mathbf{R}$, at least one industry, $i \in \mathbf{I}$, with $r_i = r$, then:

Theorem 1 *For any economy $\mathbf{E} = (\mathbf{R}, \mathbf{I}, \rho, n)$ satisfying the Hierarchy Principle and any $\beta \in (0, 1)$,*
(i) If population units in $\mathbf{E}$ satisfy the Rank-Size Rule with scale factor α and exponent β, then industrial locations satisfy the NAS Rule with scale factor $\alpha/(1-\beta)$ and exponent β.
(ii) Conversely, if $\mathbf{I}$ is a full set of industries for $\mathbf{E}$, and if industrial locations satisfy the NAS Rule with scale factor α and exponent β, then population units satisfy the Rank-Size Rule with scale factor $\alpha(1-\beta)$ and exponent β.

Proof: (i) First suppose that population units satisfy the Rank-Size Rule with scale factor α and exponent $\beta \in (0, 1)$. Then by the Hierarchy Principle together with (17) it follows that for each industry $i \in \mathbf{I}$,

$$\begin{aligned} \overline{R}_i &= \frac{1}{n_i} \int_0^{r_i} \rho(x) dx = \frac{1}{n_i} \int_0^{r_i} \alpha x^{-\beta} dx \\ &= \frac{\alpha}{n_i} \left[\frac{1}{1-\beta} x^{1-\beta} \right]_0^{r_i} = \frac{\alpha}{1-\beta} n_i^{-\beta} \end{aligned} \tag{19}$$

and hence that industrial locations satisfy the NAS Rule with scale factor $\alpha/(1-\beta)$ and exponent β.

(ii) Next suppose that $\mathbf{I}$ is a full set of industries for $\mathbf{E}$ and that industrial locations satisfy the NAS Rule with scale factor α and exponent $\beta \in (0, 1)$. Then for each $r \in \mathbf{R}$ there is some industry $i \in \mathbf{I}$

with $r = r_i = n_i$, so that by the Hierarchy Principle together with (16) and (17) we obtain the following identity for all $r \in \mathbf{R}$

$$\frac{1}{r}\int_0^r \rho(x)dx = \overline{R} = \alpha r^{-\beta} \;\Rightarrow\; \int_0^r \rho(x)dx = \alpha r^{1-\beta} \tag{20}$$

Finally, by differentiating this identity in r it follows that

$$\rho(r) = \alpha(1-\beta)r^{-\beta} \tag{21}$$

and hence that population units satisfy the Rank-Size Rule with scale factor $\alpha(1-\beta)$ and exponent β. ■

The key feature of this proof is its obvious simplicity. In addition, it allows both the Rank-Size Rule and NAS Rule to be expressed as exact power functions. Finally it shows that this equivalence is only meaningful for exponents in the open unit interval, $\beta \in (0,1)$ [as exhibited for example by Japan]. In particular it reveals the fundamental differences between $\beta < 1$, $\beta = 1$, and $\beta > 1$. For the "classic" Rank-Size Rule, $\beta = 1$, the integral in (19) becomes infinite and right hand side of (20) reduces to a constant. The behavior of $\beta > 1$ is seen from (19) and (21) to be even worse in terms of its implied relation between the Rank-Size Rule and NAS Rule.

But while this proof is quite enlightening, the continuous model itself is plagued by conceptual difficulties. There is of course the well-known problem of interpreting a continuum of population units (i.e., interpreting ρ as a density). But even more serious is the fact that the choice of rank 0 for the largest population unit implies from (18) that the population $\rho(0)$ must be *infinite*. One could of course modify the above model by choosing a positive rank value, r_0, for the highest order population unit. However, an analysis of this case shows that the simplicity of the above relations all but disappear.[27] So we choose to leave this continuous model as is.

A more realistic discrete approach is developed below which overcomes many of these problems, but requires more a more delicate analysis. In addition this discrete approach requires weaker formulations of the basic Rank-Size Rule and NAS Rule which focus on the right tail of "sufficiently small" population units. Hence it is our belief that the continuous formulation above reveals in the simplest possible way the essential relationship between the NAS Rule, Hierarchy Principle, and Rank-Size Rule.

6.2 A discrete model

In the following development we shall alter the above terminology in a manner more suitable for analysis. For purposes of this section we now take the set of population units (MEAs) to be the set of *positive*

[27] Note however that for both the classic case, $\beta = 1$, and the case $\beta > 1$, a positive value of positive value of r_0 does have the advantage of removing the infinite-integral problem. But neither of these cases yield a log-linear expression for $\overline{R}$ in terms of n. Moreover, for $\beta > 1$ the resulting expression is not even monotone decreasing in n. Hence the desired relationship still fails to hold in these cases.

integers, $\mathbf{R} = \{1, 2, ..\}$, and replace the above continuum of population units with a *ranked population sequence*, $N = (N_r : r \in \mathbf{R})$, consisting of positive population values subject to the ranking: $N_r \geq N_s$ whenever $r < s$. Hence the largest population has rank 1 and so on. There are of course only finitely many population units in any real world setting, so that actual populations can be interpreted as *initial segments* $(N_1, .., N_n)$ of ranked population sequences. For every population sequence there is a uniquely associated *upper-average sequence*, $\overline{N} = (\overline{N}_r : r \in \mathbf{R})$, defined for all $r \in \mathbf{R}$ by

$$\overline{N}_r = \frac{1}{r} \sum\nolimits_{s=1}^{r} N_s \tag{22}$$

Since averages of a monotonically nonincreasing sequence are also monotonically nonincreasing, it follows that $\overline{N}$ is also a ranked population sequence. More importantly, note that expression (22) is precisely the discrete analogue of expression (17) above, and hence that role of the Hierarchy Principle in the present context is to focus interest on the connections between these two ranked population sequences.

The corresponding discrete analogues of the Rank-Size and NAS Rules assert that the r^{th} terms of N and $\overline{N}$, respectively, are negative power functions of r. However, should be clear from (22) that N_r and $\overline{N}_r$ *cannot* simultaneously be power functions of r (since sums of power functions are not power functions). But the continuous case above suggests that some connection of this type ought to be possible. The key idea here is to replace power functions with the weaker notion of "asymptotic power laws":

Definition 1 *For any ranked population sequence, $N = (N_r : r \in \mathbf{R})$, and exponent, $\beta > 0$,*
*(i) N is said to satisfy an **asymptotic β-power law** with scale factor $\alpha > 0$, iff the sequence of ratios*

$$\theta_r = \frac{N_r}{\alpha r^{-\beta}} \ , \ r \in \mathbf{R} \tag{23}$$

converge to unity, i.e., iff

$$\lim_{r \to \infty} \theta_r = 1 \tag{24}$$

(ii) If in addition it is true that

$$\lim_{r \to \infty} r \left[\theta_r - \left(\frac{r}{r-1} \right)^{\beta} \theta_{r-1} \right] = -\beta \tag{25}$$

*then N is said to satisfy a **strong asymptotic β-power law** with scale factor α*

Hence N satisfies an asymptotic β-power law iff its terms can be written as an approximate power function

$$N_r = \alpha r^{-\beta} \theta_r \tag{26}$$

with the approximation improving as r becomes large (i.e., as population units become small). The motivation for the stronger condition (25) is much less clear. This condition is analyzed in Appendix 8.3,

where it is shown that condition (25) always holds for the limiting case in which all error factors θ_r are identically one. Moreover, it is shown that for all other sequences (θ_r) converging to one, this condition guarantees that convergence will be of order $1/r$ (i.e., that the sequence must converge to one is at least as fast as the sequence $1/r$ converges to zero). This strong convergence property highlights one key difference between asymptotic power laws for ranked population sequences N and their upper-average sequences $\overline{N}$. Since averaging is by nature a smoothing operation, it is reasonable to expect that any type of convergence for N should imply even faster convergence for $\overline{N}$. This intuition is confirmed by the following key result (proved in Appendix 8.3):

Theorem 2 *For any exponent, $\beta \in (0,1)$, a ranked population sequence N satisfies an asymptotic β-power law with scale factor α iff its upper-average sequence $\overline{N}$ satisfies a strong asymptotic β-power law with scale factor $\alpha/(1-\beta)$.*

To restate this result in a manner paralleling the continuous case above, it is convenient to make a number of simplifying assumptions. First observe that if two or more population units are of exactly the same size then population "rank" is somewhat ambiguous (particularly with respect to the Hierarchy Principle). So to avoid unnecessary complications we here model populations by a *strictly ranked population sequence*, $N = (N_r : r \in \mathbf{R})$, with population ranking: $r < s \Leftrightarrow N_r > N_s$. Next, we focus on economies with a *full* set of industries, $\mathbf{I}$, so that if $\mathbf{R}_i \subseteq \mathbf{R}$ denotes the set of population units occupied by industry i, then for each population unit r there is some $i \in \mathbf{I}$ with $r \in \mathbf{R}_i \subseteq \{1, 2, .., r\}$. For purposes of analysis, it is enough to focus on a single representative industry of this type, which we now denote as industry r. Hence associated with N is a unique *industrial-location sequence*, $\mathbf{L} = (\mathbf{R}_r : r \in \mathbf{R})$ with $r \in \mathbf{R}_r \subseteq \{1, 2, .., r\}$ for each $r \in \mathbf{R}$. For this representative economy, $\mathbf{E} = (N, \mathbf{L})$, we now say that population units satisfy a *Rank-Size Rule* with scale factor $\alpha > 0$ and exponent $\beta > 0$ iff the population sizes in N are of the form (26) with ratio sequence (θ_r) satisfying (24). Similarly, if for each set of industrial locations, $\mathbf{R}_r$, we let $n_r = \#\mathbf{R}_r$, so that the average population size of these locations is denoted by

$$\overline{N}_r = \frac{1}{n_r} \sum\nolimits_{s \in \mathbf{R}_r} N_s \tag{27}$$

then we now say that industrial locations in economy $\mathbf{E}$ satisfy a *Number-Average Size (NAS) Rule* with scale factor $\alpha > 0$ and exponent $\beta > 0$ iff for all industries $r \in \mathbf{R}$ the average population sizes $\overline{N}_r$ are of the form

$$\overline{N}_r = \alpha n_r^{-\beta} \theta_r \tag{28}$$

with ratio sequence (θ_r) satisfying (24). If in addition this ratio sequence satisfies the limit condition (25) then industrial locations in $\mathbf{E}$ are said to satisfy a *strong NAS law*. Finally, industries in economy $\mathbf{E}$

are said to satisfy the *Hierarchy Principle* iff the sets of industrial locations for all industries $r \in \mathbf{R}$ are given by the corresponding r^{th} initial segments of $\mathbf{R}$, i.e., iff

$$\mathbf{R}_r = \{1, 2, .., r\} \tag{29}$$

With these definitions, the desired equivalence result for the discrete case is given by the following corollary to Theorem 2:

Corollary 1 *For any economy* $\mathbf{E} = (N, \mathbf{L})$ *satisfying the Hierarchy Principle and any exponent* $\beta \in (0, 1)$*, population units in* $\mathbf{E}$ *satisfy a Rank-Size Rule with scale parameter* $\alpha > 0$ *and exponent* β *iff industrial locations in* $\mathbf{E}$ *satisfy a strong NAS Rule with scale parameter* $\alpha/(1-\beta)$ *and exponent* β*.*

Proof: We need only observe from expression (27) that in the presence of the Hierarchy Principle, the sequence of average population sizes (29) is precisely the upper-average sequence, $\overline{N}$, for the (strictly) ranked population sequence N. Hence the result is an immediate consequence of Theorem 2 with the Rank-Size Rule and NAS Rule treated as instances of asymptotic power laws. ■

6.3 Empirical evidence

For 1995/96, Figure 9 shows plots of (a) average size versus the number of industry-choice MEAs for each industry, (b) average size of the largest MEAs up to each rank, and (c) size versus rank of MEAs. Recall that plot (b) gives the upper bound for plot (a), and notice that (in a manner similar to Figure 4 above) it is almost log linear. Moreover, for industries with the number of industry-choice MEAs greater than 10 (i.e., for a sufficiently large r relative to rank 1), the average size of the industry-choice MEAs is almost coincident with its upper bound, as expected from Corollary 1. Hence during this period, Japanese industries appear to be quite consistent with the NAS Rule.

This Corollary also asserts that in the presence of the Hierarchy Principle (Figure 6), log linearity of plot (b) could be very well be a consequence of log linearity for the MEA rank-size distribution in plot (c), even though log linearity in (c) is not nearly as strong as in (b). As stated in the Corollary, this discrepancy could be due to the fact that convergence to log linearity for plot (b) is expected to be *faster* than for plot (c). However, it could of course also be due to a number of other factors. As one important possibility here, notice that failure of log linearity in plot (c) is most evident for small MEAs with populations less than 300,000. However, there is likely to be some misspecification of MEAs in this range. In particular, only those counties possessing a "densely inhabited district" with population greater than 50,000 are qualified to be listed as MEAs. So this could be a source of error in the log-linear regression.

Note also that there is a clear difference between the slopes of the log-linear regression for plot (b) [$\sim$ .7] and plot (c)[$\sim$ .99]. Again this disagreement is heavily influenced by the values for the smaller MEAs, and hence may in part be due to the misspecification of these MEAs. Also, note that while plot (b) and plot (c) must coincide at rank = 1, plot (c) should lie strictly below plot (b) for ranks ≥ 2 . Thus, fitting a log-linear curve to rank-size distribution (c) would likely produce the estimated slope steeper than that for the upper-average plot (b). Yet another source of error here may be the Hierarchy Principle itself, which is seen from Figure 6, to hold only approximately for Japan.

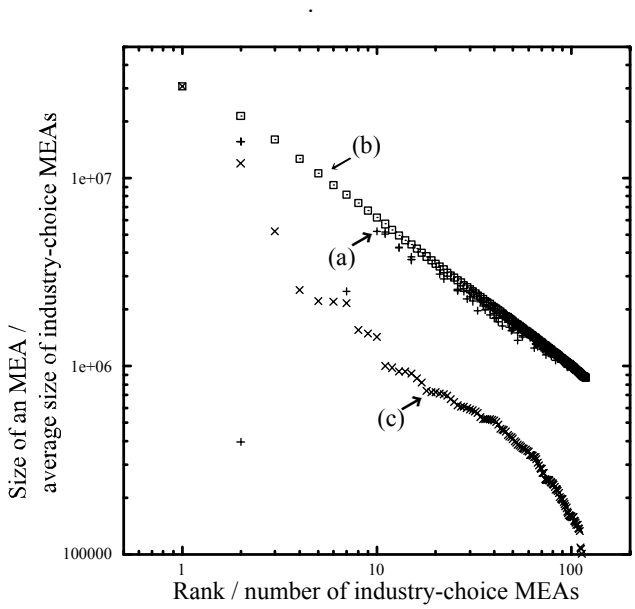

Figure 9: The Rank-Size Rule and the average size of industry-choice MEAs (1995/96)

7 Discussion and conclusions

In this paper, we have proposed a new approach to the analysis of industrial location patterns that focuses on the presence or absence of industries rather than their percentage distributions across locations. In particular we have found that by choosing MEAs as the appropriate geographic unit, this approach reveals a strong empirical relation between the average size and number of MEAs occupied by industries, designated here as the NAS Rule. This rule appears to have certain implications for regional development policies as well as for spatial economic theory in general. We address these issues in turn, and close with a brief discussion of certain directions for further research.

7.1 Some implications for regional development

An important regional policy objective is to identify those industries that are potentially sustainable in a given region (MEA). Our analysis suggests that such sustainability depends not only on region-specific factors,[28] but also the global structure of the regional system. Such global consideration are often ignored in regional industrial policy decisions. For example, several peripheral cities in Japan recently attempted to attract new IT industries to boost their economy, motivated mainly by the fact that these industries are currently growing the fastest. But IT industries are by nature *high-order* industries, i.e., are typically found in large MEAs. Our findings suggest that there may in fact be little freedom in the location pattern of industries, and in particular that there is a stable relationship between the number and size of MEAs in which a given industry can successfully locate. With respect to size in particular, the Hierarchy Principle suggests that there is the *critical MEA size* for each industry, and that only MEAs with sizes greater than this level can provide viable locations for that industry. Hence to attract such high-order industries, it would appear that these cities should focus first on attracting those lower-order (feasible) industries that will most help to stimulate regional growth. More generally, knowledge of prevailing global industrial location patterns should enhance the efficiency of regional industrial development policies.

7.2 Some implications for regional economic theory

One implication of the NAS Rule for regional economic theory is to underscore the need for more spatially disaggregated models. Most theoretical analyses of economic location are based on simple, highly aggregated models. Perhaps the popular spatial simplification is the widely used two-region model. For example, many models of economic agglomeration have been developed within this setting.[29] Such models are often a good starting point for analysis in that they tend to allow the simplest possible formalizations of spatial economic behavior. However, when it comes to explaining the actual spatial distribution of industries, say at the national level, such models are generally not very useful. In particular, they are generally of little help in formulating practical regional industrial development policies in that they do not allow explicit quantitative evaluations of the potential viability of given industries in given cities or regions. In addition, these simplified models often yield drastically different spatial equilibrium configurations (e.g., complete concentration of industries in one region, or complete uniformity of industries across regions) depending on parameter choices. But our empirical results suggest that in reality there may be greater spatial regularity than is implied by these models. Indeed it would appear that too much

[28] An example is the so-called first nature advantage of certain regions, such as the presence of natural harbors or oil fields. Regional factor endowment can also be a crucial determinant of industrial sustainability when production factors are immobile across regions, which is likely the case for the international and short-term analyses.

[29] See, e.g., Fujita and Thisse [15] for a survey.

aggregation of both regions and industries can sometimes obscure the actual regularities embodied in industrial location patterns. Our results thus suggest that a certain degree of model complexity may (paradoxically) be essential to reveal the simple structure of actual spatial economies.

A second implication relates to the theoretical relations established between the Rank-Size Rule, the Hierarchy Principle, and the NAS Rule (Theorem 1 and Corollary 1). In particular, the first two of these regularities have both been studied extensively from theoretical viewpoints. Efforts to account theoretically for the Rank-Size Rule go back at least to the work of Simon [36], who derived this rule as the steady state of a simple random city-growth process. Certain more economically based steady-state approaches have also been developed, as for example in Duranton [9] and Gabaix [17]. With respect to the Hierarchy Principle, theoretical interest has focused mainly on industrial production and demand externalities. In a variety of spatial modeling contexts, such externalities can be shown to lead to the formation of industrial agglomerations in equilibrium, as for example in the "lock-in effect" of industrial locations suggested by Fujita, Krugman, and Mori [16]. But to our knowledge, these studies have thus far been quite independent. Hence the theoretical results established here suggest a possible connection between these two lines of research. In particular, by combining the models developed for explaining the Rank-Size Rule and the Hierarchy Principle, it should be possible to give behavioral explanations for the NAS Rule as well. Such investigations will be pursued in subsequent work.

7.3 Additional directions for further research

There are at least two possible extensions of the research presented here that should be mentioned. First, our analysis has focused on global industrial location patterns at the national level. But, in the same way that "administrative cities" can be questioned as the relevant geographic unit, one may ask whether countries are indeed the relevant geographic collection of units. While there is no question that national boundaries are generally sharp barriers in terms of both population and industrial location, it nevertheless may often be true that certain subsets of regions within a given country represent more meaningful spatial economic systems than the entire country itself. For example, in the same way that commuting patterns are used to defined MEAs, one may take the patterns of interregional travel behavior to define relatively cohesive subsystems of regions within national economies. (In Japan for example, it appears that based on interregional travel flows there is a natural nesting of three regional systems: "Tokyo"$\supset$"Osaka"$\supset$"Nagoya", as discussed in Mori and Nishikimi [28]). From this viewpoint, it is then natural to ask whether empirical regularities of industrial location patterns are more readily identified at these subsystem levels. In particular, are these regularities the same as those at the national level, or

are they qualitatively different? [For Japan in particular, it turns out that the NAS Rule holds and is roughly the same for the "Tokyo", "Osaka", and "Nagoya" regions, but is not evident in smaller regional subsystems.] Such questions will be pursued further in subsequent research.

Second, while the present analysis has focused on comparisons of location patterns between industries, there may in fact be a wide variation in the locational patterns of functional units within firms (e.g., headquarters, research and development, manufacturing plant, etc.). In addition, these location patterns may in fact be similar across industries (see for example Duranton and Puga [12]). In fact, it has been pointed out that there is often a positive correlation between the size of a city and the number of cooperate control linkages emanating from that city (e.g., Fujita and Tabuchi [14]; Pred [33]; Ross [35]). It is not yet clear how this intra-firm spatial organization relates to our present findings on industrial spatial patterns. But given the central role played by multi-unit firms in modern economies, this is clearly an important direction for future research.[30]

[30]See, e.g., Duranton and Puga [12] and Ohta and Fujita [30] for the recent theoretical development for the location of multi-unit firms.

8 Appendix

8.1 Location patterns under alternative employment thresholds

Here, we plot the relationships between average size and number of industry-choice MEAs using alternative threshold employment levels to identify the presence of industries in MEAs. In particular, we now say that a given industry is present in an MEA iff the employment size of the industry in the MEA is greater than a *threshold fraction* , t, of the total population in the MEA. In the text we have used an implicit threshold fraction, $t = 0$. Figure 10 (a,b,c) shows the case for 1995/96 with $t = 0.1, 0.25$, and 0.5, respectively. Here we can observe that more industries drop from the "log- linear relationship" as the threshold fraction is increased. This is to be expected, since industries with small employment levels in particular are automatically excluded by larger threshold fractions. Note however that the majority of industries still remain around the log linear relationship established for $t = 0$.

8.2 Location patterns by industrial sectors

Figure 11 shows the relation between the average size and number of industry-choice MEAs separately for each one-digit industrial sector (manufacturing, services, wholesale and retail) in 1995/96. Manufacturing is seen to be the most diverse in terms of its location pattern, while services and wholesale-retail are mostly ubiquitous. It is remarkable that essentially the same log-linear relationship is exhibited by all three industrial sectors (except for the two outliers in manufacturing).

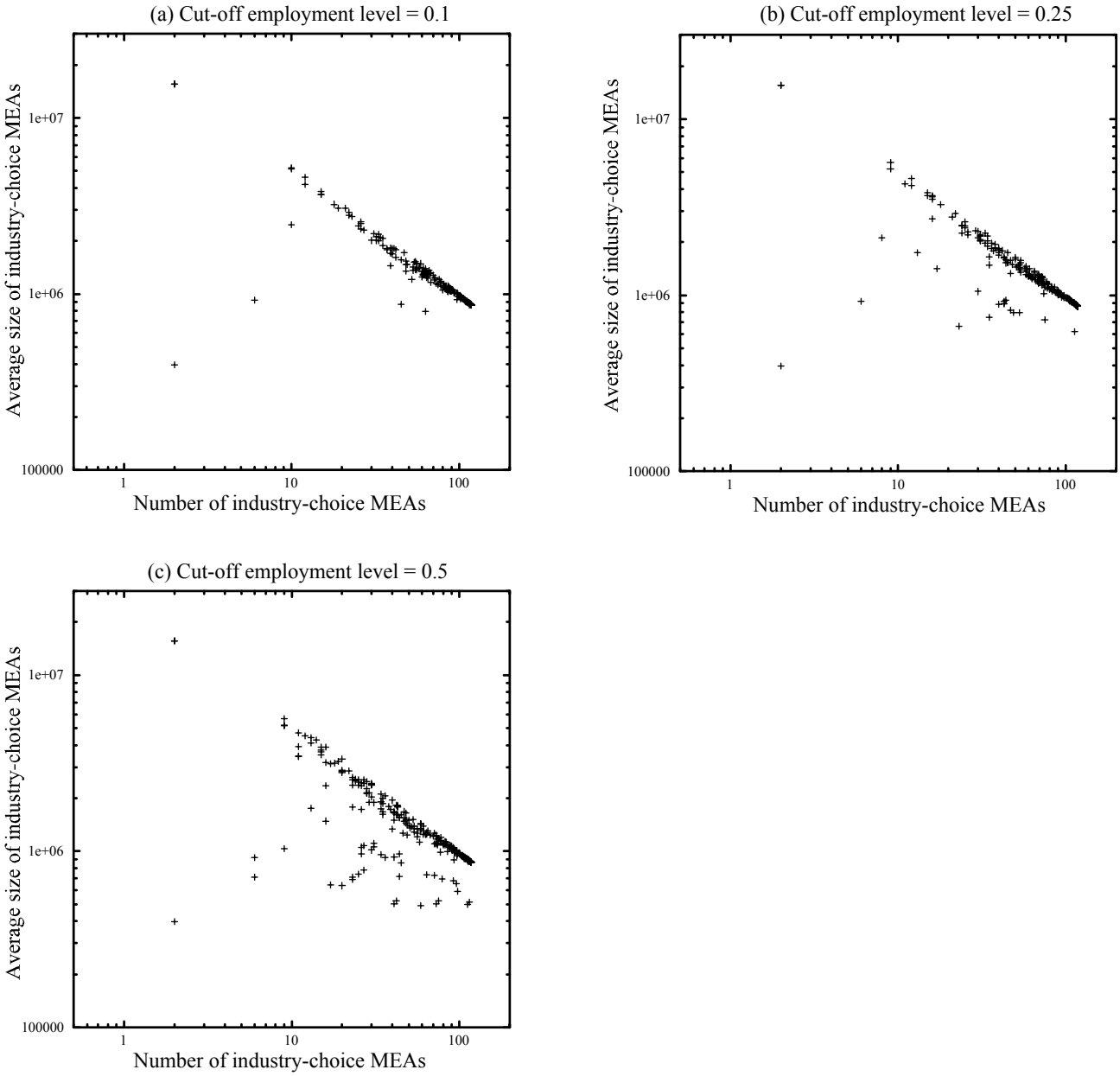

Figure 10: Location pattern under different cut-offs for employment size (1995/96)

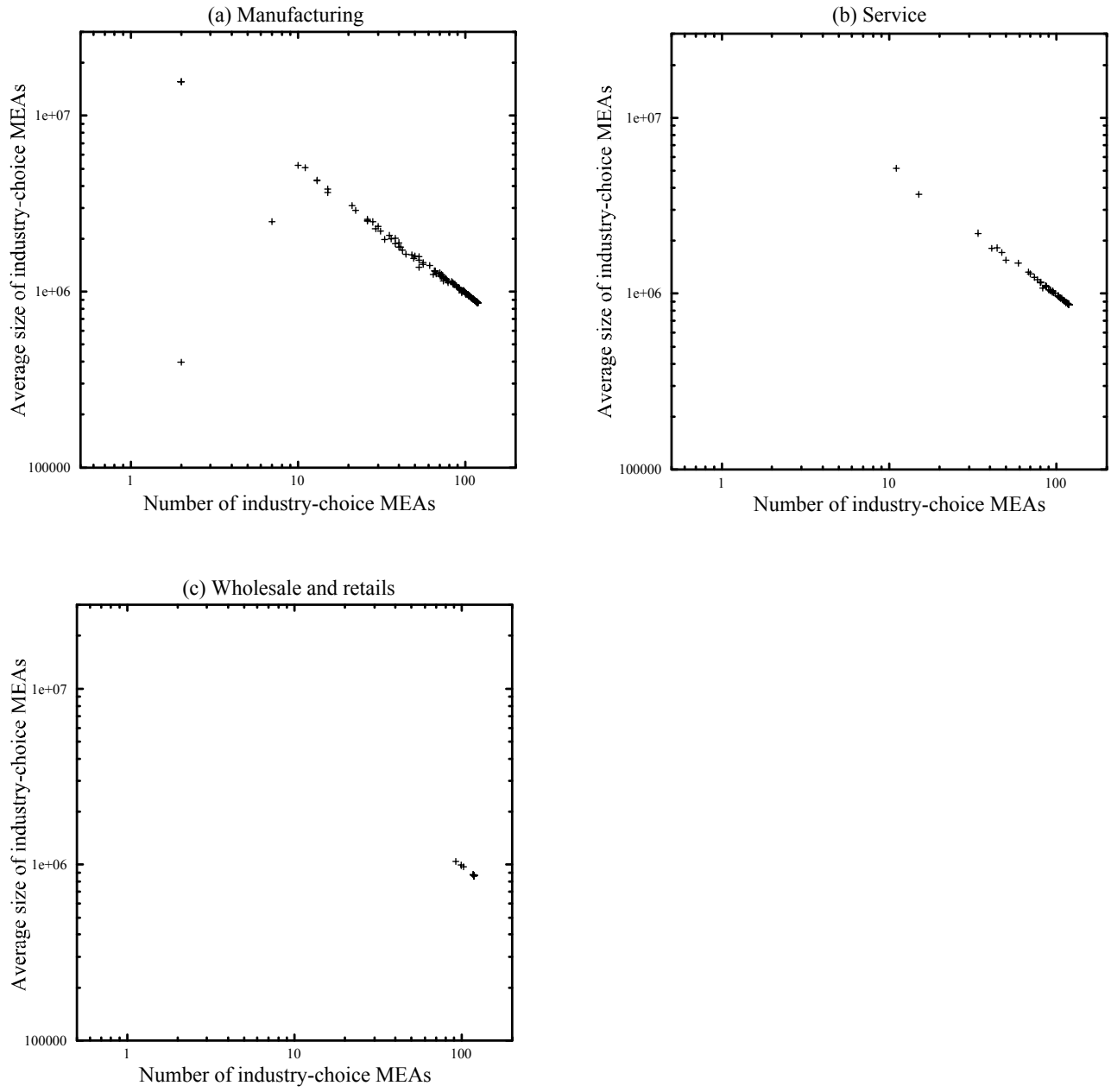

Figure 11: Average size of industry-choice MEAs by sectors (1995/96)

8.3 Formal Analysis of the Discrete Model

The main objective of this section is to prove Theorem 2 in the text. To do so, we begin with a number of preliminary results. The first two are designed to illuminate the properties of condition (25) for strong asymptotic β-power laws. We begin by showing that this condition holds identically for the sequence $(\theta_r \equiv 1)$, i.e., that

Proposition 1 *For all $\beta > 0$,*

$$\lim_{r\to\infty} r\left[1 - \left(\frac{r}{r-1}\right)^{\beta}\right] = -\beta \tag{30}$$

Proof: *If the (continuously differentiable) functions f and g are defined respectively for all $x > 0$ by*

$$f(x) = 1 - \left(\frac{x}{x-1}\right)^{\beta} \tag{31}$$

$$g(x) = 1/x \tag{32}$$

so that $\lim_{x\to\infty} f(x) = 0 = \lim_{x\to\infty} g(x)$, then it follows by L'Hospital's Rule that

$$\lim_{x\to\infty} x\left[1 - \left(\frac{x}{x-1}\right)^{\beta}\right] = \lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} \tag{33}$$

must hold whenever the last limit exists. But

$$f'(x) = -\beta\left[\frac{1}{x-1} - \frac{x}{(x-1)^2}\right] = \frac{\beta}{(x-1)^2} \tag{34}$$

$$g'(x) = -\frac{1}{x^2} \tag{35}$$

then imply that

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = -\beta \lim_{x\to\infty} \left(\frac{x}{x-1}\right)^2 = -\beta \tag{36}$$

and hence that (30) must hold. ■

Next we show that for all convergent sequences, $(\theta_r) \to 1$, condition (25) implies that this convergence must be at least of order $1/r$:

Proposition 2 For any sequence (θ_r) with $\lim_{r\to\infty} \theta_r = 1$, if *(25)* holds for some $\beta > 0$ then

$$|\theta_r - \theta_{r-1}| = O(1/r) \tag{37}$$

Proof: First note that for all $r > 1$,

$$\begin{aligned} |\theta_r - \theta_{r-1}| &= \left| \left[\theta_r - \left(\frac{r}{r-1} \right)^{\beta} \theta_{r-1} \right] + \left[\left(\frac{r}{r-1} \right)^{\beta} \theta_{r-1} - \theta_{r-1} \right] \right| \\ &\leq \left| \theta_r - \left(\frac{r}{r-1} \right)^{\beta} \theta_{r-1} \right| + \left| \left(\frac{r}{r-1} \right)^{\beta} \theta_{r-1} - \theta_{r-1} \right| \\ &= \left| \theta_r - \left(\frac{r}{r-1} \right)^{\beta} \theta_{r-1} \right| + |\theta_{r-1}| \left[\left(\frac{r}{r-1} \right)^{\beta} - 1 \right] \end{aligned} \tag{38}$$

But since this in turn implies that

$$\frac{|\theta_r - \theta_{r-1}|}{1/r} = r \left| \theta_r - \left(\frac{r}{r-1} \right)^{\beta} \theta_{r-1} \right| + |\theta_{r-1}| \left| r \left[1 - \left(\frac{r}{r-1} \right)^{\beta} \right] \right| \tag{39}$$

it follows from the limit condition $\lim_{r\to\infty} \theta_r = 1$ together with Proposition 1 above that

$$\lim_{r\to\infty} \frac{|\theta_r - \theta_{r-1}|}{1/r} = (\beta) + (1)(\beta) = 2\beta \tag{40}$$

Hence convergence is of order $1/r$ and the result is established. ■

The next two results are directed toward the proof of Theorem 2. To motivate the first result, observe from (19) in the text (with $\alpha = 1$, $\beta \in (0,1)$ and $r = r_i = n_i$) that

$$\frac{1}{r} \int_0^r x^{-\beta} dx = \frac{r^{-\beta}}{1-\beta} \;\Rightarrow\; \frac{1}{r^{1-\beta}} \int_0^r x^{-\beta} dx = \frac{1}{1-\beta} \tag{41}$$

$$\Rightarrow \lim_{r\to\infty} \frac{1}{r^{1-\beta}} \int_0^r x^{-\beta} dx = \frac{1}{1-\beta} \tag{42}$$

The following result establishes a limiting discrete analogue of this result:

Lemma 1 For all $\beta \in (0,1)$,

$$\lim_{r\to\infty} \frac{1}{r^{1-\beta}} \sum\nolimits_{s=1}^{r} s^{-\beta} = \frac{1}{1-\beta} \tag{43}$$

Proof: The result follows from standard upper and lower integral approximations to the summation in (43), namely:

$$\int_1^{r+1} x^{-\beta} dx \leq \sum\nolimits_{s=1}^{r} s^{-\beta} \leq \int_0^r x^{-\beta} dx \tag{44}$$

By calculating these integrals we see that

$$\begin{aligned} &\frac{1}{1-\beta} \left[(r+1)^{1-\beta} - 1 \right] \leq \sum\nolimits_{s=1}^{r} s^{-\beta} \leq \frac{1}{1-\beta} r^{1-\beta} \\ &\Rightarrow \frac{1}{1-\beta} \left[(\frac{r+1}{r})^{1-\beta} - \frac{1}{r^{1-\beta}} \right] \leq \frac{1}{r^{1-\beta}} \sum\nolimits_{s=1}^{r} s^{-\beta} \leq \frac{1}{1-\beta} \end{aligned} \tag{45}$$

Hence by taking limits,

$$\frac{1}{1-\beta} = \frac{1}{1-\beta}\lim_{r\to\infty}\left[\left(\frac{r+1}{r}\right)^{1-\beta} - \frac{1}{r^{1-\beta}}\right] \leq \lim_{r\to\infty}\frac{1}{r^{1-\beta}}\sum\nolimits_{s=1}^{r} s^{-\beta} \leq \frac{1}{1-\beta}$$

we obtain the desired result. ■

Next we extend this results to include error factors converging to one:

Lemma 2 For all $\beta \in (0,1)$ and sequences (θ_r) with $\lim_{r\to\infty}\theta_r = 1$,

$$\lim_{r\to\infty}\frac{1}{r^{1-\beta}}\sum\nolimits_{s=1}^{r} s^{-\beta}\theta_s = \frac{1}{1-\beta} \tag{46}$$

Proof: If $\lim_{r\to\infty}\theta_r = 1$ then for each $\varepsilon > 0$ there must be some r_ε sufficiently large to ensure that $\theta_r < 1+\varepsilon$ for all $r \geq r_\varepsilon$ **.** Hence for each $r > r_\varepsilon$,

$$\begin{aligned}
\sum\nolimits_{s=1}^{r} s^{-\beta}\theta_s &= \sum\nolimits_{s=1}^{r_\varepsilon-1} s^{-\beta}\theta_s + \sum\nolimits_{s=r_\varepsilon}^{r} s^{-\beta}\theta_s \\
&\leq \sum\nolimits_{s=1}^{r_\varepsilon-1} s^{-\beta}\theta_s + (1+\varepsilon)\sum\nolimits_{s=r_\varepsilon}^{r} s^{-\beta} \\
&\leq \sum\nolimits_{s=1}^{r_\varepsilon-1} s^{-\beta}\theta_s + (1+\varepsilon)\sum\nolimits_{s=1}^{r} s^{-\beta} \\
\Rightarrow\ \frac{1}{r^{1-\beta}}\sum\nolimits_{s=1}^{r} s^{-\beta}\theta_s &\leq \frac{1}{r^{1-\beta}}\sum\nolimits_{s=1}^{r_\varepsilon-1} s^{-\beta}\theta_s + (1+\varepsilon)\frac{1}{r^{1-\beta}}\sum\nolimits_{s=1}^{r} s^{-\beta}
\end{aligned} \tag{47}$$

Hence by taking limits in (47) we see from Lemma 1 above that:

$$\lim_{r\to\infty}\frac{1}{r^{1-\beta}}\beta > 1\theta_s \leq (0) + (1+\varepsilon)\frac{1}{1-\beta} \tag{48}$$

The same argument with r_ε large enough to ensure that $\theta_r > 1-\varepsilon$ for all $r \geq r_\varepsilon$ is easily seen to imply that

$$\lim_{r\to\infty}\frac{1}{r^{1-\beta}}\sum\nolimits_{s=1}^{r} s^{-\beta}\theta_s \geq (1-\varepsilon)\frac{1}{1-\beta} \tag{49}$$

Finally, since $\varepsilon > 0$ was chosen arbitrarily, the result follows from (48) and (49). ■

With these preliminary results, we can now establish:

Theorem 2. *For any exponent, $\beta \in (0,1)$, a ranked population sequence N satisfies an asymptotic β-power law with scale factor α iff its upper-average sequence $\overline{N}$ satisfies a strong asymptotic β-power law with scale factor $\alpha/(1-\beta)$.*

Proof: To establish this result we first observe that if for any choice of positive scalars α and $\overline{\alpha}$ we define the ratio sequences (θ_r) and $(\overline{\theta}_r)$ for all $r \in \mathbf{R}$ by

$$\theta_r = \frac{N_r}{\alpha r^{-\beta}} \tag{50}$$

$$\overline{\theta}_r = \frac{\overline{N}_r}{\overline{\alpha} r^{-\beta}} \tag{51}$$

then without loss of generality we may rewrite (22) as

$$\overline{\alpha} r^{-\beta} \overline{\theta}_r = \frac{1}{r} \sum\nolimits_{s=1}^{r} \alpha s^{-\beta} \theta_s \ , \ r \in \mathbf{R} \tag{52}$$

In these terms, it suffices to show that if the scale factors are chosen to satisfy $\overline{\alpha} = \alpha/(1-\beta)$, then the ratio sequence (θ_r) satisfies (24) iff the ratio sequence $(\overline{\theta}_r)$ satisfies (24) and (25). We begin by showing that if $\overline{\alpha} = \alpha/(1-\beta)$ then

$$\lim_{r \to \infty} \theta_r = 1 \ \ \Rightarrow \ \ \lim_{r \to \infty} \overline{\theta}_r = 1 \tag{53}$$

To do so, simply observe from (52) that

$$\overline{\theta}_r = \frac{\alpha}{\overline{\alpha}} \frac{1}{r^{1-\beta}} \sum\nolimits_{s=1}^{r} s^{-\beta} \theta_s \tag{54}$$

and hence from Lemma 2 that

$$\lim_{r \to \infty} \theta_r = 1 \ \ \Rightarrow \ \ \lim_{r \to \infty} \overline{\theta}_r = \frac{\alpha}{\overline{\alpha}} \left(\frac{1}{1-\beta} \right) = 1 \tag{55}$$

Next we consider the sequence of first differences implied by (52). In particular, if one multiplies through (52) by r to obtain

$$\sum\nolimits_{s=1}^{r} \alpha s^{-\beta} \theta_s = \overline{\alpha} r^{1-\beta} \overline{\theta}_r \tag{56}$$

then subtracting the same expression evaluated at $r-1$ yields the following sequence of relations:

$$\alpha r^{-\beta} \theta_r = \overline{\alpha} r^{1-\beta} \overline{\theta}_r - \overline{\alpha} (r-1)^{1-\beta} \overline{\theta}_{r-1} \tag{57}$$

Next, dividing through by $\overline{\alpha} r^{-\beta}$ to obtain:

$$\begin{aligned} \frac{\alpha}{\overline{\alpha}} \theta_r &= r \overline{\theta}_r - \left(\frac{r-1}{r} \right)^{-\beta} (r-1) \overline{\theta}_{r-1} \\ &= \overline{\theta}_r + (r-1) \left[\overline{\theta}_r - \left(\frac{r-1}{r} \right)^{-\beta} \overline{\theta}_{r-1} \right] \end{aligned} \tag{58}$$

It follows that if we again set $\overline{\alpha} = \alpha/(1-\beta)$ then by (53) and (58),

$$\begin{aligned} &\lim_{r \to \infty} \theta_r = 1 \ \ \Rightarrow \ \ \lim_{r \to \infty} \overline{\theta}_r = 1 \\ &\Rightarrow \ \ 1 - \beta = 1 + \lim_{r \to \infty} \left\{ (r-1) \left[\overline{\theta}_r - \left(\frac{r-1}{r} \right)^{-\beta} \overline{\theta}_{r-1} \right] \right\} \\ &\Rightarrow \ \ -\beta = \lim_{r \to \infty} \left(\frac{r-1}{r} \right) \left\{ r \left[\overline{\theta}_r - \left(\frac{r-1}{r} \right)^{-\beta} \overline{\theta}_{r-1} \right] \right\} \\ &\Rightarrow \ \ -\beta = \lim_{r \to \infty} \left[\overline{\theta}_r - \left(\frac{r-1}{r} \right)^{-\beta} \overline{\theta}_{r-1} \right] \end{aligned} \tag{59}$$

Thus that the the ratio sequence $(\overline{\theta}_r)$ also satisfies (25), and it may be concluded that $\overline{N}$ satisfies a strong asymptotic β-power law with scale factor $\alpha/(1-\beta)$ whenever N satisfies an asymptotic β-power law with scale factor α. Conversely, if the ratio sequence $(\overline{\theta}_r)$ satisfies (24) and (25) with scale factor $\overline{\alpha} = \alpha/(1-\beta)$, then again by (58):

$$\begin{aligned}(1-\beta)\theta_r &= \overline{\theta}_r + (r-1)\left[\overline{\theta}_r - \left(\frac{r-1}{r}\right)^{-\beta}\overline{\theta}_{r-1}\right] \\ &= \overline{\theta}_r + \left(\frac{r-1}{r}\right)\left\{r\left[\overline{\theta}_r - \left(\frac{r-1}{r}\right)^{-\beta}\overline{\theta}_{r-1}\right]\right\}\end{aligned}$$

$$\begin{aligned}\Rightarrow \quad (1-\beta)\lim_{r\to\infty}\theta_r &= \lim_{r\to\infty}\overline{\theta}_r + \lim_{r\to\infty}\left(\frac{r-1}{r}\right)\left\{r\left[\overline{\theta}_r - \left(\frac{r-1}{r}\right)^{-\beta}\overline{\theta}_{r-1}\right]\right\} \\ &= 1 + \lim_{r\to\infty}\left\{r\left[\overline{\theta}_r - \left(\frac{r-1}{r}\right)^{-\beta}\overline{\theta}_{r-1}\right]\right\} \\ &= 1-\beta\end{aligned}$$

$$\Rightarrow \quad \lim_{r\to\infty}\theta_r = 1 \tag{60}$$

Hence the ratio sequence (θ_r) satisfies (24) , and we may conclude that N satisfies an asymptotic β-power law with scale factor α whenever $\overline{N}$ satisfies a strong asymptotic β-power law with scale factor $\alpha/(1-\beta)$. ■

Finally, it is of some interest to note that for $\beta > 1$, our definitions do yield certain trivial log linear relationships. For example if N satisfies an exact β-power law, say $N_r = \alpha r^{-\beta}$ with $\beta > 1$ then by noting convergence of the infinite sum:

$$\sum\nolimits_{s=1}^{\infty} s^{-\beta} = c_\beta > 0 \tag{61}$$

it follows trivially that the sequence of error factors (θ_r) defined by

$$\theta_r \equiv \frac{\overline{N}_r}{\alpha c_\beta r^{-1}} = \frac{r^{-1}\sum_{s=1}^{r}\alpha s^{-\beta}}{\alpha c_\beta r^{-1}} = \frac{1}{c_\beta}\sum\nolimits_{s=1}^{r} s^{-\beta} \tag{62}$$

must converge to one. Hence if N satisfies an exact β-power law then its upper average sequence $\overline{N}$ must always satisfy an asymptotic b-power law with $b = 1$. But since the critical slope coefficient is identically one, such relations appear to have little substance.[31]

[31] However, such relations do imply that if the upper average sequence $\overline{N}$ were to satisfy any β-power law (asymptotic or exact) with $\beta > 1$, then the underlying ranked population sequence N could never satisfy a rank-size rule (asymptotic or exact).

References

[1] Auerbach, F., "Das gesetz der bevölkerungskonzentration," *Petermanns Geographische Mitteilungen* 59, 73-76 (1913).

[2] Berry, B., *Geography of Market Centers and Retail Distribution*, Prentice Hall (1968).

[3] Black, D. and Henderson, V., "Urban evolution in the USA," mimeograph, Brown University (1998).

[4] Carroll, G., "National city-size distributions: what do we know after 67 years of research?," *Progress in Human Geography* 6, 1-43 (1982).

[5] Christaller, W., *Die Zentralen Orte in Suddeutschland*, Gustav Fischer (1933), English translation by C.W. Baskin, *Central Places in Southern Germany*, Prentice Hall (1966).

[6] Dicken, P. and Lloyd, P.E., *Location in Space: Theoretical Perspectives in Economic Geography*, 3rd ed., New York: HarperCollins (1990).

[7] Dobkins, L. and Ioannides, Y., "Dynamic evolution of the size distribution of U.S. cities," in J.M. Huriot and J.-F. Thisse (eds.), *Economics of Cities*, Cambridge: Cambridge University Press, 217-260 (2000).

[8] Dumais, G., Ellison, G., Glaeser, E.L., "Geographic concentration as a dynamic process," *Review of Economics and Statistics* 84(2), 193-204 (2002).

[9] Duranton, G., "City size distribution as a consequence of the growth process," mimeograph, London School of Economics (2002).

[10] Duranton, G., and Overman H.G., "Testing for localization using micro-geographic data," mimeograph, Londons School of Economics (2002).

[11] Duranton, G. and Puga, D., "Diversity and specialization in cities: why, where and when does it matter?" *Urban Studies* 37, 535-555 (2000).

[12] Duranton, G. and Puga, D., "From sectoral to functional urban specialization," mimeograph, London School of Economics (2001).

[13] Ellison, G. and Glaeser, E.L., "Geographic concentration in U.S. manufacturing industries: a dartboard approach," *Journal of Political Economy* 105, 889-927 (1997).

[14] Fujita, M. and Tabuchi, T., "Regional growth in postwar Japan," *Regional Science and Urban Economics* 27, 643-670 (1997).

[15] Fujita, M. and Thisse, J.-F., *Economics of Agglomeration*, forthcoming, Cambridge: Cambridge University Press, Ch.2 (2001).

[16] Fujita, M., Krugman, P., and Mori, T., "On the evolution of hierarchical urban systems," *European Economic Review* 43, 209-251 (1999).

[17] Gabaix, X., "Zipf's law for cities: an explanation", *The Quarterly Journal of Economics* 114, 738-767 (1999).

[18] Henderson, V., *Urban Development*, Oxford: Oxford University Press (1988).

[19] Henderson, V., "Medium size cities," *Regional Science and Urban Economics* 27, 583-612 (1997).

[20] Henderson, V., Kuncoro, A. and Turner, M., "Industrial development in cities," *Journal of Political Economy* 103, 1067-1090 (1995).

[21] Ioannides, Y.M. and Overman, H.G., "Spatial evolution of the US urban system," Discussion paper No.0018, Tufts University (2000).

[22] Isard, W., *Location and Space-Economy*, Cambridge, MA: MIT Press (1956).

[23] Kanemoto, Y. and Tokuoka, K., "The proposal for the standard definition of the metropolitan area in Japan," Discussion paper No.J-55, Center for International Research on the Japanese Economy, Tokyo University, Tokyo, Japan (2001).

[24] Krugman, P., *Geography and Trade*, Cambridge, MA: MIT Press (1991).

[25] Lösch, A., *The Economics of Location*, Jena, Germany: Fischer (1940), English translation, New Haven, CT: Yale University Press (1954).

[26] Marshall, J., *The Structure of Urban Systems*, Toronto, University of Toronto Press (1989).

[27] Mori, T. and Nishikimi, K., "A Note on the index for industrial localization and diversity in cities," in progress.

[28] Mori T. and Nishikimi, K., "Recursive pattern formation in the industry and population location," in progress.

[29] Office of Management and Budget, "Standards for defining metropolitan and micropolitan statistical areas," Federal Register Vol.65, No.249 (2000).

[30] Ohta, M. and Fujita, M., "Communication technology and spatial configurations of intrafirm units and residential units," *Journal of Regional Science* 21, 87-104 (1991).

[31] Overman, H. and Ioannides, Y.M., "Cross-sectional evolution of the US city size distribution," mimeograph, London School of Economics (2000).

[32] Parr, J., "A note on the size distribution of cities over time," *Journal of Urban Economics* 18, 199-212 (1985).

[33] Pred, A., *The Spatial Dynamics of U.S. Urban-Industrial Growth*, Cambridge: MIT Press (1966).

[34] Rosen, K.T. and Resnick, M., "The size distribution of cities: an examination of the Parato law and primacy," *Journal of Urban Economics* 8, 165-186 (1980).

[35] Ross, S., *The Urban System and Networks of Corporate Control*, Greenwich, CT: JAI Press (1991).

[36] Simon, H., "On a class of skew distribution functions," *Biometrika* 42, 425-440 (1955).

[37] Zipf, G.K., *Human Behavior and the Principle of Least Effort: an Introduction to Human Ecology*, Cambridge, Massachusetts: Addison Wesley (1949).