TOMOYA MORI
Kyoto University

TONY E. SMITH
University of Pennsylvania

A Reconsideration of the NAS Rule from an Industrial Agglomeration Perspective

Japan experienced rapid urbanization after the World War II as indicated, for example, by the fact that the population share of Densely Inhabited Districts (DID), nearly doubled between 1950 and 2000, from 34.9 percent to 65.2 percent, while accounting for only 3.3 percent of the national land.[1] Moreover, this rapid urbanization does not appear to be a simple proportional increase of economic activities in all urban areas. Rather, the spatial distributions of industries and population within the 258 metro areas (cities) of Japan are quite skewed. The population of the largest city, Tokyo, exceeded 30 million in 2000 and accounted for more than a quarter of the national population. The ten largest cities together accounted for more than a half of the national population. Moreover, if the *industrial diversity* of a given city is defined in terms of the number of industries exhibiting significant agglomeration within that city (see the section on cluster-based choice cities and industries below), then the population sizes of cities also appear to be highly correlated with their industrial diversities (see the section on the hierarchy principle).

We especially thank Gary Burtless, Gilles Duranton, Vernon Henderson, Wen-Tai Hsu, Yannis Ioannides, Edwin Mills, Art O'Sullivan, Janet Pack, John Quigley, and Ping Wang for stimulating discussions. This research has been partially supported by The Murata Science Foundation, The Grant in Aid for Research (Nos. 17330052, 18203016, 19330049) of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

1. *Densely Inhabited Districts* are defined in the Population Census of Japan (Japan Statistics Bureau 1980, 2000) as a geographic areas having a residential population of at least 5,000 with a population density greater than 4,000/km^2. These DIDs are also used in the World Urbanization Prospects (United Nations 2007) definition of "urban shares" of population.

A similarly drastic urbanization has been observed, for example, in South Korea where the "urban share of population" nearly quadrupled from 21.4 percent in 1950 to 79.6 percent in 2000. The corresponding percentages for the United States, western Europe, and western Asia are from 64.2 percent to 79.1 percent, 63.8 percent to 75.3 percent, and 16.5 percent to 40.4 percent, respectively.

Against this background, our main interest is to ask whether these skewed spatial distributions of industries and population exhibit any clear relationship, or whether they might simply have happened by chance. In Mori, Nishikimi, and Smith (2008), a strong empirical regularity was identified between the size and industrial composition of cities in Japan. This regularity, designated as the Number-Average Size (NAS) Rule, asserts a negative log-linear relation between the number and average population size of those cities where a given industry is present, that is, of the *choice cities* for that industry. More recently, this same regularity (with comparable definitions of industries and cities) has been reported for the United States by Hsu (2008).

But despite the strong empirical regularity of the NAS Rule, there still remains the statistical question of whether such location patterns could simply have occurred by chance. Of particular importance here is the focus of this rule on the presence or absence of industries in each city, rather than on the percentage distribution of industries across cities. Indeed, chance occurrences of certain choice cities may be quite likely if, for example, one includes cities where only a single industrial establishment happens to appear. Hence there is a need to clarify exactly what constitutes a *substantial industrial presence* in a given city. Although it is possible to characterize *substantial* in terms of some threshold number or share of industrial establishments or employment, such conventions are necessarily ad hoc. Hence an alternative approach is proposed in a companion paper, Mori and Smith (2009b), which characterizes *substantial* in terms of significant industrial agglomeration. More specifically, this approach utilizes the statistical procedure developed in Mori and Smith (2009a) to identify spatially explicit patterns of significant clustering (agglomeration) for each industry. In this context, the desired choice cities for each industry are taken to be those (economic) cities that share at least a part of a significant spatial cluster for the industry and therefore are designated as *cluster-based choice cities*.

With this new definition, it is shown in Mori and Smith (2009b) that the NAS Rule not only continues to hold for Japan but in some ways is even stronger. In particular, the few industrial outliers identified for the NAS Rule in the original analysis of Mori, Nishikimi, and Smith (2008) are shown here to be *without exception* industries for which no significant spatial agglomerations can be identified. Hence these results serve to suggest that the NAS Rule may in fact be an observable consequence of underlying coordination between spatial agglomerations of industry and population.

But unlike the original analysis in Mori, Nishikimi, and Smith (2008), the NAS Rule in Mori and Smith (2009b) is examined only for 2001. To that end, there remains the question of whether this rule continues to exhibit the same persistence over time that was seen in the original analysis. The results for 1981 have now been completed, and indeed they confirm persistence of the NAS Rule over this twenty-year period.[2] Thus the main objective of the present paper is to develop these new results and to compare them with the original analysis in Mori, Nishikimi, and Smith (2008).

This persistence is particularly remarkable given that it does not arise from simple proportional growth, such as a proportional increase in the number of cluster-based choice cities across industries or a proportional increase in the average sizes of these cities. On the contrary, there has been a substantial churning of these choice cities across industries (as developed later in the section on churning of cluster-based choice cities and industries).

It was also shown in Mori, Nishikimi, and Smith (2008) that there is an intimate theoretical connection between the NAS Rule and both the classical Rank Size Rule for cities and Christaller's (1966) Hierarchy Principle for industrial location behavior. Thus a final objective of the present paper is to analyze the persistence of these two additional regularities with respect to cluster-based industry-choice cities over the given twenty-year period.

To develop these results, we begin in the next section with an overview of both the city and industry data employed in this analysis. The third section, after the data section, then focuses on cluster-based choice cities (and cluster-based choice industries), as constructed in Mori and Smith (2009b). These cities are analyzed with respect to their relative employment concentrations and their key churning properties with respect to industry mix. This is followed in the subsequent section with a review of the NAS Rule itself and a presentation of the new findings of persistence. In the same section, this pattern of persistence is extended to the Hierarchy Principle and Rank Size Rule. Finally, the paper concludes with a brief discussion of ongoing work and directions for further research.

Data

The data used in the present analysis is very similar to that used in the original two-period analysis of Mori, Nishikimi, and Smith (2008). In the discussion below,

2. This involved restricting the admissible set of industries to those that are comparable between the two time periods. It should also be mentioned here that several months of computer time were required to generate sufficient random cluster patterns for testing the significance of agglomerations in each of these observed industrial location patterns.

Figure 1. Municipality Boundaries, 2001

Source: National Statistics Center of Japan (2009).

we focus on the differences between the two. We begin with city data and then consider industrial data.

Cities

The basic regional units in the present study that are used to identify economic cities and industrial agglomerations are municipalities. The 3,230 municipalities used in Mori, Nishikimi, and Smith (2008) were based on data in 2000. In the present paper, the municipality boundaries in 2000 are converted to the latest definition used in 2001, which creates certain minor differences. More important, the 13 major municipalities (such as Tokyo, Osaka, Nagoya, and Kyoto) have been divided into their individual *wards*, which are comparable in size to most other municipalities. This increases the total to 3,363 as of October 1, 2001. Finally, since industrial agglomerations are identified in terms of road-network distances (see the industry clusters section), we focus on the 3,207 municipalities that are geographically connected to the major islands of Japan (that is, Honshu, Hokkaido, Kyushu, and Shikoku) via a road network (refer to figure 1). The excluded municipalities account for only 1.6 percent of the total population in both 1980 and 2000 and should not have a significant effect on the analysis.

In terms of these basic regional units, an (*economic*) *city* is formally defined to be an Urban Employment Area (UEA), as proposed originally by Kanemoto and Tokuoka (2002). Each UEA is designed to be an urban area of Japan that is comparable to the metropolitan areas (MAs) of a Core Based Statistical Area (CBSA) in the United States.[3] Hence each UEA consists of a core set of municipalities designated as its business district (BD) together with a set of suburban municipalities from which workers commute toward the BD. Following Kanemoto and Tokuoka (2002), UEAs are constructed as aggregations of municipalities by a recursive procedure that is detailed in Mori, Nishikimi, and Smith (2008). Basically this construction starts with a large "seed" municipality, designated as the central municipality of the UEA. This in turn is extended to a BD and an appropropriate set of suburban municipalities. However, the analysis in Mori, Nishikimi, and Smith (2008) used only Metropolitan Employment Areas (MEAs), that is, UEAs with central municpality populations of at least 50,000. In the present analysis, we include all UEAs as defined by Kanemoto and Tokuoka (2002) that have central municipality populations of at least 10,000. Those with central munipality populations below 50,000 are designated as Micropolitan Employment Areas. This broader definition yields 309 cities (UEAs) in 1980 and 258 cities in 2000 (compared with the respectively smaller sets of 105 and 113 MEAs used in Mori, Nishikimi, and Smith 2008).

Industries

As in Mori, Nishikimi, and Smith (2008), the industrial 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 2001 (Japan Statistics Bureau 1981, 2001) and are applied to the respective population data in 1980 and 2000. But unlike in Mori, Nishikimi, and Smith (2008), the present analysis focuses on manufacturing. For while services and wholesale-retail industries tend to be found almost everywhere, that is, they are ubiquitous, manufacturing industries exhibit a much larger diversity of location patterns at the three-digit level. Therefore, as observed in Mori, Nishikimi, and Smith (2008), the NAS Rule itself is far more interesting for manufacturing industries.[4]

3. For the definition of a CBSA, see U.S. Office of Management and Budget (2000).

4. This can be seen quite dramatically in figure B1 of appendix B in Mori, Nishikimi, and Smith (2008), where the 125 manufacturing industries shown (see footnote 14) are a subset of those used here.

There were 152 and 164 manufacturing industries at the three-digit level in 1981 and 2001, respectively. Hence, to achieve comparability between industrial location patterns in these two years, industries in each year have been aggregated in a manner that yields the largest number of common classifications with a positive number of establishments for both years. This aggregation resulted in 147 common manufacturing industrial classifications for both years.[5] This number is further reduced to 139 industries that exhibit at least some degree of significant agglomeration (as discussed in industry clusters below).

Cluster-Based Choice Cities and Industries

As stated in the introduction, the central objective of this paper is to reexamine the NAS Rule with respect to cluster-based choice (cb-choice) cities for each industry. Since the identification of cb-choice cities for industries is developed fully in Mori and Smith (2009a, 2009b), we only sketch the main ideas below. Given the definition of cities above, the focus here will be on the identification of significant industrial clusters. These clusters are used to define cb-choice cities for each industry when we discuss the definitions of cluster-based choice cities and industries. From the city perspective, there is a completely parallel concept of cb-choice industries for each city. This is followed with a brief consideration of the relative industrial employment concentration in cb-choice cities relative to all other cities. Finally, the churning of industrial locations is considered from both industry and city viewpoints in the last subsection.

Industrial Clusters

Our approach to the identification of significant clusters of regions (municipalities) for a given industry is closely related to the statistical clustering procedures proposed by Besag and Newell (1991), Kulldorff and Nagarwalla (1995), and Kulldorff (1997). To test for the presence of clusters, these procedures start by postulating an appropriate null hypothesis of no clustering. In the present case, this hypothesis is characterized by a uniform distribution of industrial locations across regions.[6] Such clustering procedures then seek to determine the single most significant cluster of regions with respect to this hypothesis. Candidate clusters are typically defined to be approximately cir-

5. See appendix A for the details of this industrial aggregation.

6. Here *uniformity* is defined with respect to an areal measure of the economic area of each region (municipality). Details of this measurement procedure are given in Mori and Smith (2009a).

cular areas containing all regions having centroids within some specified distance of a given reference point (such as the centroid of a central region). The approach developed in Mori and Smith (2009a) extends these procedures in two ways. First, the notion of a circular cluster of regions is extended to the more general (metric-space) concept of a *convex solid*, as defined with respect to the shortest travel-distance metric on the given set of regions.[7] Next, individual (convex solid) clusters, C, are extended to the more global concept of cluster schemes. If the set of relevant regions (municipalities), r, is denoted by R, then each cluster scheme, $\mathbf{C} = (R_0, C_1, \ldots C_{k_\mathbf{C}})$, is taken to be a partition of R into one or more disjoint clusters, $C_1, \ldots C_{k_\mathbf{C}}$, together with the residual set, R_0, of all noncluster regions. Each cluster scheme then induces a family of possible location probability models, called cluster probability models, $p_\mathbf{C} = [p_\mathbf{C}(j) : j = 1, \ldots, k_\mathbf{C}]$, in which it is implicitly hypothesized that industrial establishments are more likely to be located in one of the cluster regions than in a noncluster region (and where $p_\mathbf{C}(R_0) = 1 - \Sigma_j p_\mathbf{C}(j)$). Each cluster probability model, $p_\mathbf{C}$, thus amounts formally to multinomial sampling models on its underlying cluster schemes, $\mathbf{C}$.[8]

In this context, the local problem of finding a single, most likely cluster is replaced by the global problem of finding a cluster probability model that best fits the full set of industry locations. In turn, this is seen to be an instance of the general statistical goodness-of-fit problem, that is, the problem of selecting a best-fit model from among a family of candidate probability models for a given set of sample data. While many model-selection criteria have been proposed for doing so, the criterion chosen here is the Bayesian Information Criterion (BIC) first proposed by Schwarz (1978).[9] Essentially this criterion involves a trade-off between the likelihood of the given sample data under each candidate model and the number of parameters (cluster probabilities) used in the model.[10]

To find a best cluster model with respect to this criterion, it would of course be ideal to compare all possible cluster schemes that can be constructed from the given system of regions. But even for modest numbers of regions, this is a

7. Here *shortest travel distance* is defined with respect to *road-network distance*, as detailed in Mori and Smith (2009a).

8. Other models of this type include the model-based clustering approach of Dasgupta and Raftery (1998), and the Bayesian approach of Gangnon and Clayton (2000, 2004). See Mori and Smith (2009a, footnote 7) for further discussion.

9. Among the many other model-selection criteria that are applicable here, the most prominent are Akaike's (1973) Information Criterion (AIC) and the Normalized Maximum Likelihood (NML) Criterion by Kontkanen and Myllymäki (2005). A comparison of these criteria in the present context will be presented in Smith and Mori (2009).

10. Further details are given in Mori and Smith (2009a, 2009b).

practical impossibility. Hence the approach taken in Mori and Smith (2009a) is to develop a heurisitic algorithm that searches among the set of candidate models for the best model with respect to the BIC criterion. To do so, one starts by finding the best cluster probability model with an underlying cluster scheme consisting of exactly one single-region cluster (municipality). More elaborate cluster schemes are then grown by adding new disjoint clusters or by either expanding or combining existing clusters until no further improvement in the BIC model-selection criterion is possible. The final result is thus guaranteed to yield at least a locally best cluster scheme with respect to this criterion.[11] If the set of manufacturing industries is denoted by I, then let the best cluster scheme found for industry, $i \in I$, be denoted by $\mathbf{C}_i = (R_{i0}, C_{i1}, .., C_{ik_i})$.

Cluster schemes for each of the 147 manufacturing industries were constructed for 1981 and 2001. Since the construction procedure and analysis of these cluster schemes is identical for both years, we drop time distinctions and simply take $\mathbf{C}_i$ to be a generic representation of both cluster schemes for each industry i. In addition it should be noted that both cluster schemes for each of these 147 industries contain at least one cluster, and hence are nondegenerate.

But even when cluster schemes are nondegenerate, there remains the statistical question of whether such clustering could simply have occurred by chance. Indeed, even completely random location patterns will tend to exhibit some degree of clustering.[12] Therefore, for each industry i, one can ask how the optimal criterion value, BIC_i, obtained for $\mathbf{C}_i$ compares with typical values obtained by applying the same cluster detection procedure to randomly generated spatial data. This testing procedure can be formalized in terms of the null hypothesis of complete spatial randomness, which asserts that individual establishment locations are independently and uniformly distributed over the economic areas of regions. Under this hypothesis, therefore, the probability, $P(r)$, that any given establishment will locate in region (municipality), $r \in R$, is taken to be proportional to the size of economic area of region r. Monte Carlo simulation can then be employed to estimate the sampling distribution of BIC_i under this hypothesis, and a one-sided test can be performed to determine whether the observed value of BIC_i is significantly large relative to this distribution. Those industries with clustering that is not significant at the 5 percent level are said to exhibit spurious clustering.[13]

11. See Mori and Smith (2009a) for further details.

12. In fact, the complete absence of clustering is statistically consistent with a significantly dispersed (ubiquitous) pattern of industrial locations, which is the complete opposite of clustering (agglomeration).

13. See Mori and Smith (2009a).

Among the 147 industries for which clusters were identified, all were extremely significant (with p values close to zero) except for 8 industries where complete spatial randomness could not be rejected at the 5 percent level. These include six arms-related industries (JSIC381, 383–387), together with tobacco (JSIC194), and coke manufacturing (JSIC273). Besides the small numbers of establishments in these industries, they also are rather special in other ways.[14] Tobacco manufacturing and arms-related industries are highly regulated industries, so that their location patterns are not determined by market forces. Finally, coke production is a typical declining industry in Japan (steel industries have gradually replaced coke production with less expensive powder coal after the 1970s).

Thus our present analysis is based on the remaining 139 industries that exhibit some degree of significant clustering. For these industries, the percentages of establishments included in clusters range from 39.1 percent to 100 percent (with an average of 94.1 percent) in 2001, while the corresponding percentages in 1981 range from 51.8 percent to 100 percent (with an average of 95.7 percent).[15]

Definition of Cluster-Based Choice Cities and Cluster-Based Choice Industries

For each industry with significant clustering, we can now define its set of corresponding cluster-based choice cities as follows. First, if the set of all cities (UEAs) in a given year is indexed by U, and if the subset of cities with positive employment in industry i is indexed by $U_i^+ \subseteq U$ (where again we drop time distinctions), then a city, $k \in U_i^+$, is designated as a cluster-based choice city for industry i, if and only if there is some cluster, $C_i \in \mathbf{C}_i$, such that

$$UEA_k \cap C_i \neq \varnothing. \tag{1}$$

In other words, UEA_k is a cb-choice city for i, if and only if (iff) it shares at least one positive i-employment municipality with some cluster in $\mathbf{C}_i$.[16] Let the set of cb-choice cities for i be indexed simply by U_i. To distinguish this notion from the original set of choice cities, U_i^+, proposed in Mori, Nishikimi, and Smith

14. All have less than 40 establishments, with an average of 7.89 establishments (compared with the average of 6,183 establishments for the other industries in 2001). Establishment location data are not available for tobacco manufacturing (JSIC194) in 1981 since it was operated by the national government.

15. For further discussion, see Mori and Smith (2009b).

16. Here it should be noted that the "convexity" requirement on clusters in $\mathbf{C}_i$ implies that a cluster may contain some municipalities with no i employment. Hence as a minimum condition, cb-choice cities for i are required to share cluster municipalities in $\mathbf{C}_i$ with *positive* i employment.

(2008), it is convenient to designate all cities in U_i^+ as presence-based (pb) choice cities for industry *i*.

Note that the intersection in equation 1 can be interpreted in terms of individual cities as well as industries. In particular, one may designate industry *i* as a cb-choice industry for city *k* iff $k \in U_i^+$ and equation 1 holds for some cluster, $C_i \in \mathbf{C}_i$. As a parallel to U_i, one may then index the set of cb-choice industries, $i \in I$, for city *k* by I_k. Hence, in the same way that the number ($\#U_i$) of cb-choice cities for a given industry reflects its location diversity, the number ($\#I_k$) of cb-choice industries for a given city reflects its industrial diversity. These diversity measures exhibit great variation across industries and cities alike. With respect to the 139 industries studied, $\#U_i$ ranges from 14 to 275 (with an average of 116.3) cities in 1981, and ranges from 12 to 227 (with an average of 101.7) cities in 2001. Similarly, for the 309 cities identified in 1980, $\#I_k$ ranges from 2 to 139 (with an average of 52.3) industries in 1981, and for the 258 cities identified in 2000, $\#I_k$ covers the full range from 1 to 139 (with an average of 54.8) industries. We shall examine some additional empirical properties of these dual relations in the last subsection below.

Industrial Concentration in Cluster-Based Choice Cities

Next, recall that the primary motivation for the present definition of cb-choice cities was to characterize substantial industry presence in terms of agglomeration behavior. Hence we next consider how this endogenous approach relates to more exogenous threshold approaches in terms of industrial concentration. Such concentration can be measured in terms of either employment or numbers of establishments in cities. The key finding here is that with respect to both these measures, cb-choice cities for industries do indeed exhibit larger concentrations than do other cities in which the industry is present.

To state this more precisely, let E_{ik} and N_{ik} denote respectively the employment size and number of establishments of industry *i* in city *k*. Then, the employment-concentration ratio, R_i^{emp}, of average *i*-employment in cb-choice cities, U_i, to that in all other cities with positive *i*-employment, $U_i^+ - U_i$, is given by:

$$R_i^{emp} \equiv \frac{\frac{1}{\#U_i}\sum_{k \in U_i} E_{ik}}{\frac{1}{\#U_i^+ - \#U_i}\sum_{k \in U_i^+ - U_i} E_{ik}}. \qquad (2)$$

Similarly, the establishment-concentration ratio, R_i^{est}, of the average number of i-establishments in cb-choice cities, U_i, to that in all other cities with positive i-employment, $U_i^+ - U_i$, is given by:

$$R_i^{est} \equiv \frac{\frac{1}{\#U_i}\sum_{k\in U_i} N_{ik}}{\frac{1}{\#U_i^+ - \#U_i}\sum_{k\in U_i^+ - U_i} N_{ik}}. \tag{3}$$

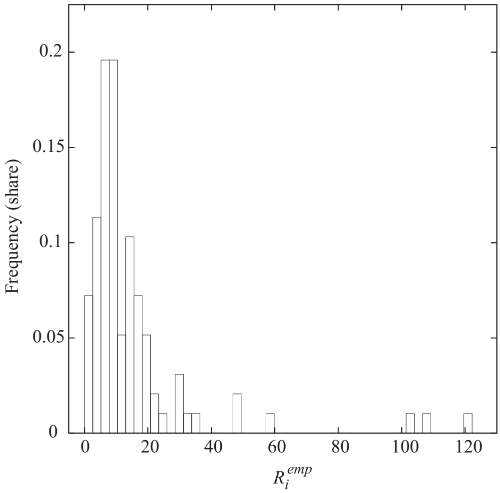
The frequency distributions of concentration ratios, R_i^{emp} and R_i^{est}, over all 139 industries, i, with significant clustering in 2001 are shown in figures 2 and 3, respectively. Here R_i^{emp} ranges from 1.17 to 121.00 (with an average value of 15.19), and R_i^{est} ranges from 1.47 to 71.74 (with an average value of 15.05). Notice in particular that *all* ratios are greater than one. Hence it is clear that cb-choice cities for each industry i do indeed exhibit relatively large concentrations without imposing ad hoc threshold sizes on such concentrations.

Churning of Cluster-Based Choice Cities and Cluster-Based Choice Industries

Recall from the section above that for each time period there is a wide range in the locational diversities of industries and the industrial diversities of cities. But even more important is the fact that there has been a considerable amount of churning of industries across cities and vice versa. One way to examine these effects is to consider changes in the number of cb-choice cities for each industry i between 1981 and 2001, as shown in figure 4. Here figure 4(a) shows these changes as calculated using the respective city boundaries identified for each year. Figure 4(b) shows these changes using the 2000 city boundaries for both years (so that only changes in industrial agglomeration patterns are reflected). In both figures, industries are ordered by their locational diversity (that is, by the number of cb-choice cities) in 1981. The vertical bar shown for each industry is divided into two segments. The length of the upper segment corresponds to the number of new cb-choice cities for this industry in 2001 that were not cb-choice cities in 1981, and the length of the lower segment is the number of old cb-choice cities in 1981 that ceased to be cb-choice cities by 2001. These two diagrams suggest that regardless of changes in city boundaries, there are significant numbers of both exiting and entering cb-choice cities for most industries.

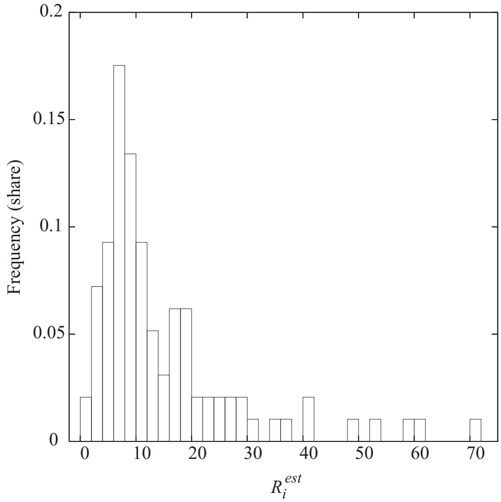
An alternative way to examine these churning effects is to measure changes in the set of cb-choice cities for each industry between these two years. If the

Figure 2. Average Employment Size of Cluster-Based Choice Cities Relative to That of Presence-based Ones



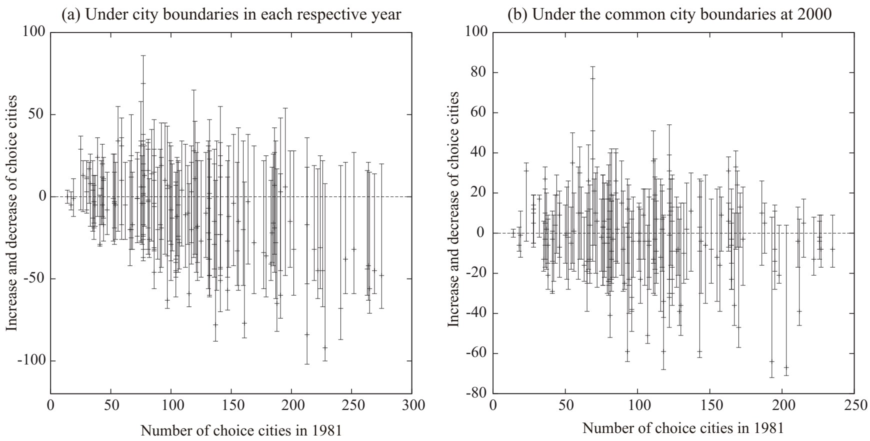
Source: Japan Statistics Bureau (2001) and authors' calculations.

Figure 3. Average Establishment Count of Cluster-Based Choice Cities Relative to That of Presence-based Ones



Source: Japan Statistics Bureau (2001) and authors' calculations.

Figure 4. Change in the Number of Choice Cities



Source: Japan Statistics Bureau (1980, 1981, 2000, 2001) and authors' calculations.

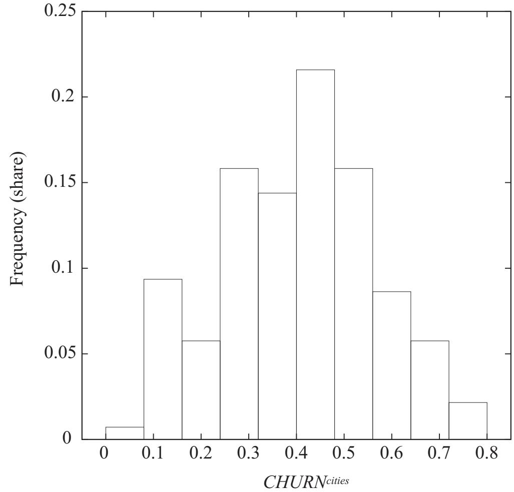
sets of cb-choice cities for each industry $i \in I$ in 1981 and 2001 (with respect to 2000 city boundaries) are denoted respectively by U_i^{1981} and U_i^{2001}, then the churning of cb-choice cities for i can be measured as follows:

(4) $$CHURN_i^{cities} = 1 - \frac{\#(U_i^{1981} \cap U_i^{2001})}{\#(U_i^{1981} \cup U_i^{2001})}.$$

Hence complete churning corresponds to $CHURN_i^{cities} = 1$, where all cb-choice cities for industry i have changed from 1981 to 2001. Similarly, $CHURN_i^{cities} = 0$ implies no churning. The frequency distribution of these churning values across all 139 industries with significant clustering is shown in figure 5. The values of $CHURN_i^{cities}$ range from 0.06 to 0.78 with an average of 0.41. Here more than half of the cb-choice cities for thirty-nine (28.1 percent) of these industries were replaced during this twenty-year period (and more than a quarter were replaced for at least 80 percent of the industries). In short, these industries have exhibited dramatic churning of their locations during this period. Similar rapid adjustments of industrial locations have been documented for France and the United States by Duranton (2007).[17]

17. The employment share–based measure of churning of industrial locations adopted by Duranton (2007) is somewhat problematic when relatively disaggregated industries are considered (as in the present study) since employment shares may often be zero.

Figure 5. Churning of Choice Cities

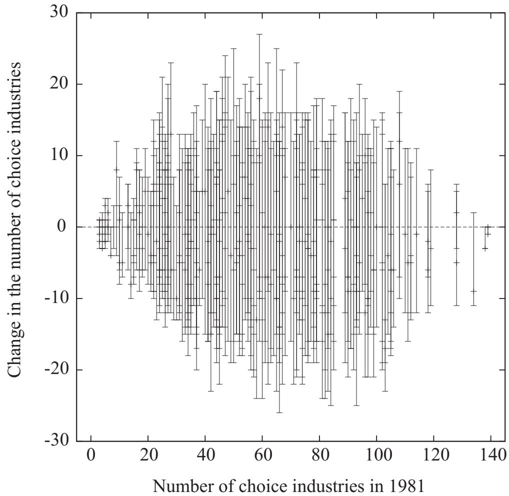


Source: Japan Statistics Bureau (1980, 1981, 2000, 2001) and authors' calculations.

Such churning can also be measured from the city perspective. Here we focus on the 258 cities identified in 2000 and use their 2000 boundaries for analysis. As a parallel to the industries analysis above, we first consider changes in the number of cb-choice industries for each city, *k*, during this twenty-year period, as shown in figure 6. Here cities are ordered on the *x*-axis in terms of their industrial diversity (that is, number of cb-choice industries) in 1981. The length of the upper segment of the vertical bar for each city *k* now corresponds to the number of new cb-choice industries for *k* in 2001 that were not cb-choice industries for *k* in 1981, and the length of the bottom segment corresponds to the number of cb-choice industries for *k* in 1981 that had ceased to be cb-choice industries for *k* by 2001.

It is clear from the figure that the change in industrial composition is smallest for the most diversified and the least diversified cities. This is partly due to the fact the industry classification is fixed, so that the number of choice industries has little room for increase in the most diversified cities. Similarly, there is little room for decrease in the least diversified cities. But, as figure 6 shows, there is also little decrease for the most diversified cities, and little increase for the least diversified cities. So the industrial diversification of cities at both ends of the spectrum appears to be relatively stable during this twenty-year period.

Figure 6. Change in the Number of Choice Industries of Cities



Source: Japan Statistics Bureau (1980, 1981, 2000, 2001) and authors' calculations.

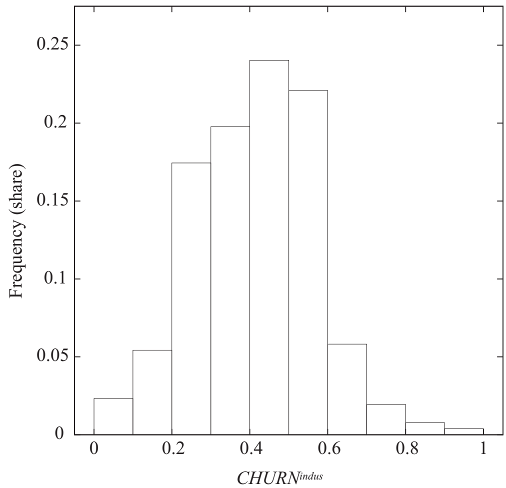
Thus, churning of cb-choice industries occurs mostly in cities with intermediate levels of industrial diversity.

As with industries, these churning effects can also be examined by measuring changes in the sets of cb-choice industries for cities. To do so, let the sets of cb-choice industries for each city *k* in 1981 and 2001 (with respect to 2001 city boundaries) be denoted by I_k^{1981} and I_k^{2001}. Then the churning of cb-choice industries for *k* can be measured as follows:

(5) $$CHURN_k^{indus} = 1 - \frac{\#(I_k^{1981} \cap I_k^{2001})}{\#(I_k^{1981} \cup I_k^{2001})},$$

where complete churning of industries for *k* again corresponds to $CHURN_k^{indus} = 1$, and where $CHURN_k^{indus} = 0$ again implies no churning. The frequency distribution of $CHURN_k^{indus}$ across all cities, *k*, is shown in figure 7. Here the values of $CHURN_k^{indus}$ take on the full range from 0.0 to 1.0, with an average of 0.58. As with cb-choice cities above, there is substantial churning of cb-choice industries. Here more than half of the cb-choice industries for 77 (30.0 percent) of these 258 cities are replaced (and more than a quarter were replaced for 216, or 83.7 percent, of these cities).

Figure 7. Churning of Choice Industries between 1981 and 2001



Source: Japan Statistics Bureau (1980, 1981, 2000, 2001) and authors' calculations.

The NAS Rule and its Associated Empirical Regularities

Given the definitions and preliminary findings above, we turn now to the major results of this paper. Here we begin with the NAS Rule itself below and consider its persistence properties for the case of Japan under our new definition of cluster-based choice cities. These persistence properties are then extended to the associated Hierarchy Principle and Rank Size Rule in the next two sections, respectively.

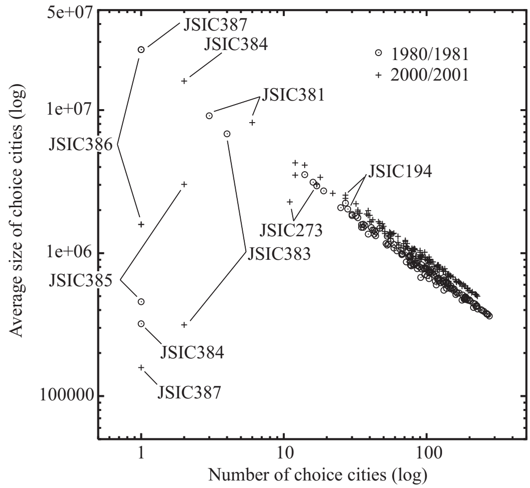
The NAS Rule

In the present setting, the Number-Average Size (NAS) Rule first formulated in Mori, Nisikimi, and Smith (2008) (in terms of pb-choice cities) now asserts that there is a log-linear relationship between the number and average size of cb-choice cities for industries. With respect to this new definition of choice cities, the main result of the present paper is shown in figure 8.

In figure 8, the logs of both the number of cb-choice cities ($\#CITY$) and average size of cb-choice cities ($\overline{SIZE}$) are plotted for the relevant 147 manufacturing industries in 1981 and 2001.[18] The specific points corresponding to the eight

18. In terms of the notation in the section on definitions of cb-choice cities and industries above, for each industry i, $\#CITY_i = \#U_i$ and $\overline{SIZE_i}$ is the average size of all cities in U_i.

Figure 8. Average Size versus the Number of Choice Cities of Industries



Source: Japan Statistics Bureau (1980, 1981, 2000, 2001) and authors' calculations.

industries with spurious clustering are indicated in figure 8, and they show that *all outliers* are in this group. Hence for those 139 industries with significant clustering, the relations shown for each year are almost exactly log linear. This can be verified by a simple OLS regression, which yields the following results for each year:[19]

$$(6)\qquad 1980-1981: \log(\overline{SIZE}) = \underset{(0.0309)}{16.92} - \underset{(0.00668)}{0.734}\ \log(\#CITY), R^2 = 0.996,$$

$$(7)\qquad 2000-2001: \log(\overline{SIZE}) = \underset{(0.0286)}{17.01} - \underset{(0.00635)}{0.716}\ \log(\#CITY), R^2 = 0.996,$$

where the values in parentheses are standard errors. It should be noted that since the dependent variables are neither normally distributed nor independent by construction, the linear estimates in equations 6 and 7 are best regarded as "curve fitting" rather than genuine statistical models (as pointed out by Eaton and Eckstein (1997, p. 452, footnote 19).[20] However, it should also be emphasized that

19. Because our city data are for 1980 and 2000 while our industry data are for 1981 and 2001, we shall sometimes denote these two periods by 1980/1981 and 2000/2001, respectively.

20. It should also be noted that if city sizes are distributed according to a power law (as implied by the Rank Size Rule below), then as pointed out by Gabaix and Ioannides (2004, section 2.2.1), the standard errors in these regressions may be grossly underestimated. But in the present case, with R^2 values almost 1, it should be clear that such standard errors add little in the way of new information.

these strong log-linear relations are not simply the result of some underlying tautology. In particular, the drastic outliers in figure 8 suggest that for industries without strong agglomeration tendencies, this NAS Rule may not be relevant at all. Hence it can be conjectured that insofar as agglomeration behavior is a reflection of economic factors, the NAS Rule is most relevant for industries where location decisions are largely governed by economic considerations.

Aside from the obvious strength of this log-linear relation, it should also be emphasized that the slopes of these two regression lines are almost the same. This can again be tested by pooling the data for both time periods, introducing a time dummy and applying standard F tests to evaluate coefficient shifts. While the statistical validity of such a test is again questionable in terms of normality and independence, the results clearly support invariance of the slope coefficient. However, the intercept does exhibit a significant shift, as can easily be seen from figure 8. So while both the numbers and average sizes of cb-choice cities for individual industries have changed, they have done so in a manner that leaves their elasticity of substitution invariant. More specifically, a 1 percent increase in the number of cb-choice cities for an industry from 1981 to 2001 corresponds roughly to a 0.7 percent decrease in the average size of these cities during the same twenty-year period.

The stability of this relation is even more remarkable in view of the dramatic churning of industries across cities during this period (as discussed in the section on churning above). In addition, there has also been a substantial reordering of city sizes themselves (as discussed below). The invariance of the NAS Rule adds further support to the conjecture that this implicit coordination between industrial and population locations is driven by the same underlying economic forces over time. Although the exact nature of these forces remains an open question, the recent model proposed by Hsu (2008) suggests that scale economies of production may constitute one important contributing factor.

The results in Hsu (2008) together with the original analysis in Mori, Nishikimi, and Smith (2008) show that this NAS Rule is intimately connected with two other well-known classical regularities of city systems, namely, Christaller's (1966) Hierarchy Principle and the Rank Size Rule of city size distributions. Hence the invariance of the NAS Rule above suggests that these two regularities may also exhibit invariance properties. We now consider each of these regularities in the context of our present manufacturing data.

The Hierarchy Principle

The Hierarchy Principle originally proposed by Christaller (1966) asserts that industries found in cities of a given size will also be found in all cities of larger sizes. The approach of Mori, Nishikimi, and Smith (2008) was to redefine this principle in terms of industrial diversity (that is, the number of choice industries for a city) rather than by population size. Hence our present version of the Hierarchy Principle asserts that industries in cities with a given level of industrial diversity (that is, a given number of cb-choice industries) will also be found in all cities with larger industrial diversities. This version is formally somewhat weaker than the original population version of Christaller and hence constitutes a necessary condition for the classical Hierarchy Principle.[21] The main advantage of this reformulation is that it allows the Hierarchy Principle to be tested without altering the industrial diversity structure of the city system. Moreover, this weaker version is in reality very closely related to the classical version. In the present case, Spearman's rank correlation between the industrial diversity levels and populations of cities is around 0.75 for both 1980/1981 and 2000/2001.

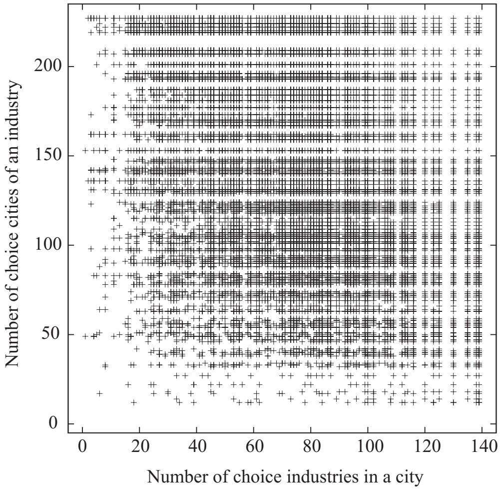
Before testing this principle in the present setting, it is useful to consider the city-industry relationships depicted graphically in figure 9, using 2001 data. Here, cities, k, are ordered by their industrial diversities (number of cb-choice industries) on the horizontal axis, and industries, i, are ordered by their locational diversities (number of cb-choice cities) on the vertical axis. A plus symbol (+) in position (k, i) indicates that k is a cb-choice city for industry i (and equivalently, that i is a cb-choice industry for city k). If we distinguish such positions as positive, then the Hierarchy Principle asserts that for each positive position (k, i), then there must also be a + in every column position $(\cdot, i)$ to the right of (k, i), indicating that all cities with industrial diversities greater than or equal to city k are also cb-choice cities for industry i. Hence it is clear from the figure that while the Hierarchy Principle does not hold perfectly, the row density of + values increases from left to right in virtually every row. Hence there is clearly a strong level of agreement with the Hierarchy Principle that could not have occurred by chance.[22]

A formal statistical test of this assertion was developed in Mori, Nishikimi, and Smith (2008). To apply this test in the present context, it suffices to outline the basic elements of the test in terms of figure 9 (see Mori, Nishikimi, and Smith 2008, section 4 for a detailed development). To do so, observe first

21. See footnote 40 in Mori, Nishikimi, and Smith (2008).

22. Note that this figure bears a strong resemblance to figure 7 in Mori, Nishikimi, and Smith (2008). The key difference for our present purposes is the new cluster-based definition of choice cites for industries. However, it should be noted that the inclusion of Micropolitan Employment Areas in the present analysis greatly expands the range of cities with small industrial diversities (at the left end of the city scale). It should also be noted that the SIC classification system for industries is by no means exact; some level of disagreement in such hierarchical relations is unavoidable.

Figure 9. Industry-Location Events



Source: Japan Statistics Bureau (2000, 2001) and authors' calculations.

that each occurrence of a full row of + values to the right of a positive position (k, i) can be regarded as a "full hierarchy event" at (k, i) in the sense that it is fully consistent with the Hierarchy Principle. However, in cases where only small fraction of + values are missing, it is natural to regard this as being closer to a full hierarchy event than if all + values were missing. To distinguish between such cases, it is appropriate to designate the fraction of positive positions to the right of each positive (k, i) as the fractional hierarchy event, H_{ki}, at (k, i). Thus $0 \leq H_{ki} \leq 1$, with $H_{ki} = 1$ denoting a full hierarchy event at (k, i). Note also that since by definition each positive position (k, i) generates a unique fractional hierarchy event (of which it is the left end point), the number, h, of fractional hierarchy events is precisely the number of positive positions (+ values) in the figure. Hence, as a measure of overall consistency with the Hierarchy Principle, we designate the average of these fractional hierarchy events as the (observed) hierarchy share,

$$p_0 = \frac{1}{h} \sum_{ki} H_{ki}, \tag{8}$$

for the given system of cities and industries. By definition, $0 \leq p_0 \leq 1$, with $p_0 = 1$ now denoting exact agreement with the Hierarchy Principle, that is, all fractional hierarchy events are full.

In this context, one possible null hypothesis for testing the Hierarchy Principle would be that this figure is the realization of a stochastic process in which

h of these + values are assigned randomly to (k, i) pairs (without replacement). However, it can be argued that this null hypothesis is too strong in the sense that it not only ignores industrial hierarchies but also ignores the basic urban structure of the city system itself. For example, major cities such as Tokyo and Osaka are implicitly treated as indistinguishable from even the smallest micropolitan cities in Japan. To preserve actual urban structure to some degree, we thus choose to hold the industrial diversity of each city fixed.[23] To test this Hierarchy Principle, the null hypothesis, H_0, adopted here is that the observed distribution of + values in figure 9 is the realization of a stochastic process that assigns random + values in a manner that preserves the industrial diversity of each city, that is, preserves the number of + values in each column of the figure. Since the industrial diversity of city k is given by the number of its cb-choice industries, $\#I_k$, it follows that this process is easily realized by randomly selecting $\#I_k$ cb-choice industries from I for each city k. By constructing a large number of such realizations, say 1,000, and calculating the hierarchy share, p_m, for each realization, $m = 1, \ldots, 1000$, one can then test the Hierarchy Principle by simply checking whether the observed hierarchy share, p_0, is unusually large relative to this sample of typical share values under H_0.

The results of this (one-sided) test in the present case provide a strong rejection of H_0 in favor of significantly large hierarchy shares. In particular, for the 2001 data in figure 9, the observed hierarchy share is $p_0 = 0.775$, while the simulated hierarchy shares under H_0 ranged from 0.622 to 0.631.[24] Thus, even when the industrial diversity structure of this city system is held fixed, the statistical evidence in favor of the Hierarchy Principle is overwhelming. A parallel application of this test to the 1981 data produced essentially the same findings, with an observed hierarchy share of $p_0 = 0.772$ and a simulated range of hierarchy shares from 0.612 to 0.618 under H_0. The similarity of these values shows that in spite of the dramatic churning of both industries and city sizes during this twenty-year period, the overall hierarchical structure of industrial locations has remained remarkably stable.

Finally, it is of interest to consider the implications of these results for the NAS Rule itself. The relation between this rule and the Hierarchy Principle is seen most easily in terms of the broader definitions of these concepts in Mori, Nishikimi, and Smith (2008), in which the classical Hierarchy Principle (in terms of city size) was used and choice cities for each industry i were taken to

23. Recall that since industrial diversity is highly rank correlated with city size, this convention tends to preserve the ordering city sizes as well.

24. The observed value of p_0 is so far above this range that larger simulation sizes would surely yield similar results.

include the larger set, U_i^+, of all cities where i is found. If this classical Hierarchy Principle were to hold exactly, and if we denote the smallest choice city for i by $\tilde{k}_i \in U_i^+$, then U_i^+ would consist precisely of all cities with populations at least as large as $\tilde{k}_i$. Moreover, $\#U_i^+$ would then be the number of cities at least as large as $\tilde{k}_i$, which is by definition the (population) rank of city $\tilde{k}_i$. Under these conditions, the NAS Rule is equivalent to a negative log-linear relation between the rank of city $\tilde{k}_i$ and the average of all city sizes at least as large as $\tilde{k}_i$, designated in Mori, Nishikimi, and Smith (2008) as the upper-average city size for $\tilde{k}_i$. If the rank and upper-average city size of each city, k, are denoted respectively by $RANK_k$ and $\widehat{SIZE}_k$ then the plot of $\log \widehat{SIZE}$ against $\log(RANK)$ for all cities in plot b of figure 10 in Mori, Nishikimi, and Smith (2008) showed a remarkably close relation to a plot of the actual values of $\log(\overline{SIZE})$ and $\log(\#CITY)$ in plot a of figure 10.[25]

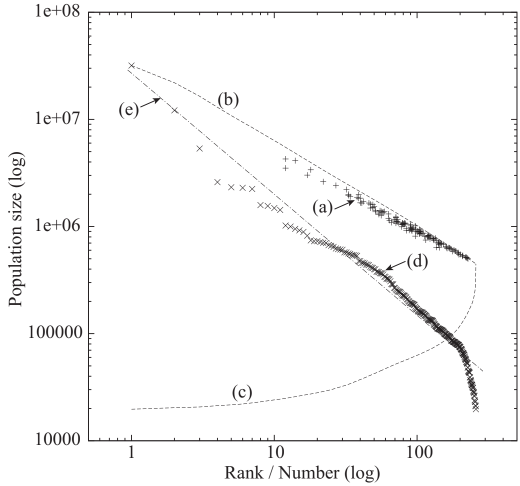
In the present context, both the definitions of choice cities (as cb-choice cities) and the Hierarchy Principle (in terms of industrial diversity) have changed. However, as noted at the beginning of this section, the rank correlation between city size and industrial diversity continues to be high. This, together with the test results above (as well as the direct evidence in figure 9), suggests that the average size of cb-choice cities in U_i ($\subset U_i^+$) should still agree reasonably well with the upper-average city size for the smallest cb-choice city, $k_i \in U_i$. This relation is demonstrated in figure 10 below, which bears a striking resemblance to figure 10 in Mori, Nishikimi, and Smith (2008). Here, using data for 2001, $\log \widehat{SIZE}$ is again plotted against $\log(RANK)$ for all cities in plot b of figure 10. Similarly, for the present definition cb-choice cities, $\log(\overline{SIZE})$ is plotted against $\log(\#CITY)$ in plot a of figure 10.[26] It is clear, therefore, that the present restriction to cb-choice cities (as well as the inclusion of Micropolitan Employment Areas) has made little difference. These relations are both very close, where the slightly flatter slope of the NAS Rule again reflects imperfections in the Hierarchy Principle.

Note also that if $\underline{SIZE}_k$ denotes the average of all city sizes smaller than city k, then in the same way that $\widehat{SIZE}_k$ represents the natural upper bound on $\overline{SIZE}_k$, the value $\underline{SIZE}_k$ represents a natural lower bound. In these terms, the plot of $\log(\underline{SIZE})$ against $\log(\#CITY)$ in plot c of figure 10 shows that within its feasible range of values, $\log(\overline{SIZE})$ is almost identical with its upper bound. As

25. As in footnote 18 above, for each industry i, $\#CITY_i$ here represents $\#U_i^+$, and $\overline{SIZE}_i$ is the average size of all cities in U_i^+.

26. Plot a is identical with the plot of 2001 data in figure 8, in which all industries with spurious clustering have now been removed.

Figure 10. City Size Distribution and the NAS Rule



Source: Japan Statistics Bureau (2000, 2001) and authors' calculations.

observed in Mori, Nishikimi, and Smith (2008), this serves to further underscore the extremely nonrandom nature of plot a.

Finally, in view of the closeness of plot a to this upper bound in plot b, it is of interest to ask whether the stability of the NAS Rule in equations 6 and 7 above is also reflected in this upper-bound relation between $\log \widehat{SIZE}$ and log(*RANK*). The corresponding regression results for 1980/1981 and 2000/2001 are shown below:

$$(9)\quad 1980-1981: \log(\widehat{SIZE}) = \underset{(0.00692)}{17.37} - \underset{(0.00143)}{0.806} \log(\#RANK), R^2 = 0.999,$$

$$(10)\quad 2000-2001: \log(\widehat{SIZE}) = \underset{(0.00718)}{17.52} - \underset{(0.00154)}{0.805} \log(\#RANK), R^2 = 0.999.$$

As with the NAS regressions in equations 6 and 7 above, an *F* test using equations 9 and 10 confirms that only the intercept has shifted in any significant way. In fact, the slope of this Hierarchy relation appears to be even more stable than the NAS Rule over this twenty-year period.

The Rank Size Rule

Finally we turn to the Rank Size Rule for systems of cities, which asserts that if all cities are ranked by population size, then the $RANK_k$ and $SIZE_k$ of cities, k, are (approximately) negatively log-linearly related, that is, that for all cities,

$$\log(SIZE) \approx \sigma + \theta \log(RANK). \quad (11)$$

The classical version of this rule also asserts that $\theta \approx -1.0$. For the present case, a plot of $\log(SIZE_k)$ against $\log(RANK_k)$ for all cities, $k \in U$, is shown in plot d of figure 10. Here again, this plot is qualitatively similar to plot c of figure 10 in Mori, Nishikimi, and Smith (2008), showing that the restriction to cb-choice cities (and inclusion of Micropolitan Employment Areas) has made little difference. Log linearity is again most evident in the central range of the plot, while the relative slopes at each end are much steeper.[27] At each extreme it appears that other socioeconomic mechanisms may be at work (as discussed further in Mori, Nishikimi, and Smith (2008)).

But our present interest focuses mainly on the relation between the Rank Size Rule in equation 11 and the NAS Rule. In Mori, Nishikimi, and Smith (2008), it was shown that in the presence of the classical Hierarchy Principle, the NAS Rule (with respect to the larger sets of choice cities U_i^+ for industries i) is asymptotically equivalent to this Rank Size Rule, that is, they satisfy the same asymptotic power law (see Mori, Nishikimi, and Smith 2008, corollary 1). Hence the above stability results for the NAS Rule in equations 6 and 7, and for the Hierarchy Principle in equations 9 and 10, suggest that stability over this twenty-year period may also be exhibited by the Rank Size Rule. Regressions of log(*SIZE*) on log(*RANK*) in periods 1980/1981 and 2000/2001 produced the following results:

$$1980-1981: \log(SIZE) = \underset{(0.0582)}{16.85} - \underset{(0.0120)}{1.094} \log(\#RANK), R^2 = 0.964, \quad (12)$$

$$2000-2001: \log(SIZE) = \underset{(0.0786)}{17.11} - \underset{(0.0168)}{1.130} \log(\#RANK), R^2 = 0.946. \quad (13)$$

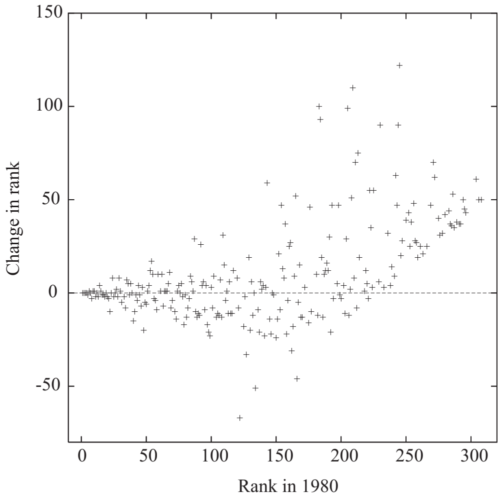
Here again an *F* test shows that only the intercept has shifted significantly and hence that the slope of this overall relation has remained fairly stable. The regression line for 2000/2001 in equation 13 is shown in plot e of figure 10.[28]

As with the NAS Rule discussed above, this stability of the Rank Size Rule is even more remarkable in view of the substantial shuffling of population ranks

27. Here, there is a slight kink at cities with populations of about 300,000 and a much sharper dip at cities below 70,000.

28. Since the data for the Rank Size relation in 1980/1981 strongly overlap the data for 2000/2001, it is difficult to show both on the same figure. Thus we choose to display only the latter.

Figure 11. Change in the Ranking of Cities Existing in Both Years



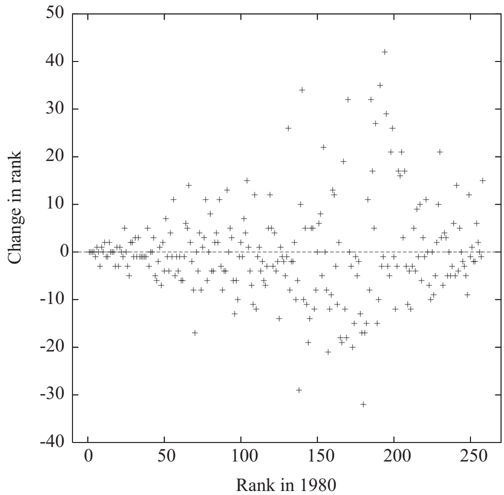
Sources: Japan Statistics Bureau (1980, 2000) and authors' calculations.

among cities. For the 246 cities that existed in both 1980 and 2000 (by our definition of cities), the changes in population ranks of these cities in 2000 are plotted against their corresponding 1980 ranks in figure 11. As is clear from the figure, while the largest cities remained relatively stable, most cities actually moved up in the rankings, with an average jump in rank of 10.25.[29] But there is also a great deal of variation in movement. For example, there are cities like Uozu with large upward jumps—from 245th (49,512) in 1980 to 123th (134,411) in 2000—and other cities like Okaya with large downward jumps—from 122th (371,850) in 1980 to 189th (435,367) in 2000.[30] In addition to this movement, there were

29. This same phenomenon (largest cities remaining stable) is observed in other countries as well. For the United States in particular, see, for example, Black and Henderson (2003).

30. The numbers in parentheses are the population levels of the city for 1980 and 2000, respectively. The most significant growths and declines of cities in the studied period seemed to be triggered by the expansion of Shinkansen (bullet train) lines. For instance, the extension of the Shinkansen line from Tokyo to Fukuoka in 1975 leads to the population growth of Fukuoka as a new center of Kyushu region from 1,762,794 to 2,323,604 (31.8 percent) and from sixth to fifth in the population rankings, while it also caused the population decline by 6.64 percent and from eighth to eleventh in the population rankings of Kitakyushu, the traditional regional center of Kyushu, located 50 km east of Fukuoka. Okayama experienced even more drastic population growth from 744,735 to 1,484,742 (99.4 percent) and moved up from fourteenth to tenth in the size ranking. There are two major reasons for this disproportionate growth. One is that Okayama became a transshipment point of the extended Shinkansen line between Tokyo and Fukuoka mentioned above. The other is the completion of the Seto-oh-hashi in 1988, the bridge connecting the main island to Shikoku island via Okayama.

Figure 12. Change in the Ranking of Cities with Boundaries Fixed at 2000



Source: Japan Statistics Bureau (1980, 2000) and authors' calculations.

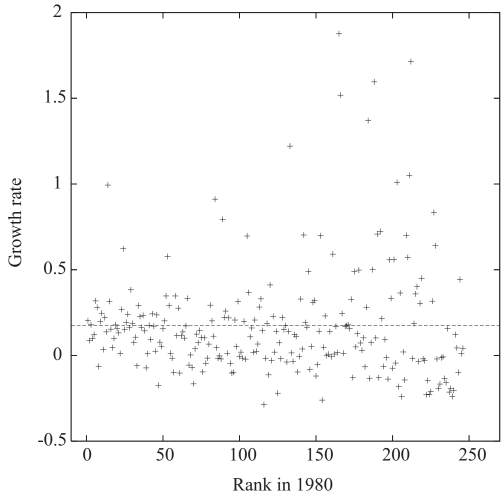
also changes in the sets of cities themselves (again with respect to our definition of cities). Most of the reduction in cities from 309 in 1980 to 258 in 2000 was due to the absorption of one city by another. In fact, most cities that exhibited large upward jumps in the rankings grew by absorption of nearby cities.[31]

One problem with these observations is that changes in the number of cities between 1980 and 2000 make it somewhat difficult to interpret the changes in rankings above. Thus even though these absolute rankings are the ones used in the regressions of equations 12 and 13, it is useful for our present purposes to consider changes in the relative rankings of the 246 cities existing in both years. This can be done by simply ranking these cities from 1 to 246 in 1980 and recording the changes of these ranks in 2000. But even here it can be argued that such changes might be due largely to the changes in city boundaries between 1980 and 2000 (resulting from our city definitions above). Hence it is useful to consider changes in these relative rankings using the city boundaries in 2000 for both years.[32] These changes in relative rankings are shown in figure 12. By construction, the average change in rankings of cities is now zero. But the range of such changes from –32 to 42 again shows wide variation (with

31. For instance, Uozu, with a population of 49,512, moved up 122 ranks by absorbing Kurobe with a population of 72,259. Similarly, Kitagami with a population of 76,633 moved up 100 ranks by absorbing Hanamaki with a population of 97,389.

32. See Overman and Ioannides (2001) for a discussion of the choice of geographical areas for cities when making intertemporal comparisons of city sizes.

Figure 13. Change in the Sizes of Cities Existing in Both Years



Source: Japan Statistics Bureau (1980, 2000) and authors' calculations.

a standard deviation of 10.1). Thus even in terms of relative rankings with fixed boundaries, the changes in rankings over this twenty-year period have been dramatic.

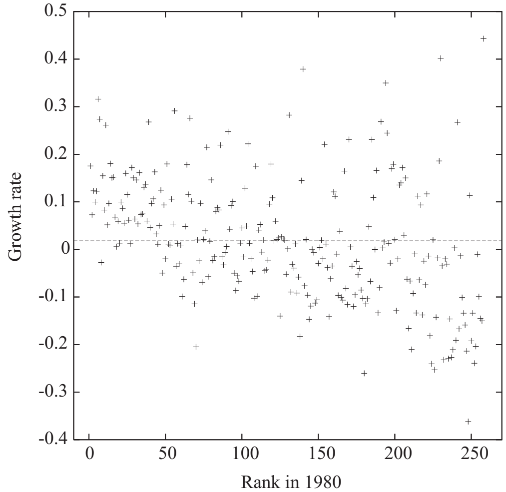
In addition to changes in rankings, there has also been very uneven growth among cities. In figure 13, the population growth rates of these 246 cities are plotted against their absolute rankings in 1980. The sizes of these cities increased by 17.5 percent on average, with a standard deviation of 32.4 percent. Note also from the figure that variation in growth rates appears to be higher for smaller cities.

In figure 14 below, growth rates are plotted using fixed city boundaries from 2000. Here the average growth rate, 1.81 percent, is now much smaller since the (usually larger) city boundaries in 2000 are used. But even here it is remarkable that the growth rates of cities range from –36.2 percent to 44.3 percent, with a standard deviation of 13.1 percent.

Finally, returning to the Rank Size Rule itself, it is of interest to compare the regression results in equations 12 and 13 with both the classical Rank Size Rule and the NAS Rule. Notice first that the overall slope of each regression is close to –1.0, and hence appears to be in rough agreement with the classical Rank Size Rule. But it is clear for the regression line shown in plot e of figure 10 for 2000–01 data that this slope is in fact a compromise between the slopes for each of the three data segments in plot d decribed above.[33] In particular, the

33. This same type of compromise is also exhibited by the regression line for 1980/1981 (not shown).

Figure 14. Change in the Sizes of Cities with Boundaries Fixed at 2000



Sources: Japan Statistics Bureau (1980, 2000) and authors' calculations.

slope of the middle range in plot d, for which log linearity is most evident, is seen to be much flatter and is indeed much closer to that of the NAS regression in equation 7 above than it is to –1.0.[34] Moreover, it can be argued that this central range is dominant in the sense that the slope of the upper average relation in plot b essentially mirrors that of this range. In fact, Mori, Nishikimi, and Smith (2008, Theorem 2) have shown that the log linearity of the rank-size distribution and that of the upper-average distribution are asymptotically equivalent with the same log slope. So, while no definitive conclusions can be drawn from such limited observations, they do suggest that the theoretical relations between the NAS Rule and both the Rank Size Rule and Hierarchy Principle developed in Mori, Nishikimi, and Smith (2008) are empirically most evident for cities in this dominant central range.

Concluding Remarks

The main purpose of this paper has been to examine the temporal stability properties of the NAS Rule under the sharper definition of cluster-based choice cities for industries proposed by Mori and Smith (2009b). In particular, it was shown above that for Japanese manufacturing industries between 1981 and

34. A regression using only this middle segment yields a slope of –0.767.

2001, the stability of this rule under presence-based choice cities continues to hold for cb-choice cities as well. In addition, it was shown that, as in Mori and Smith (2009b), similar stability properties are exhibited by the two other regularities closely related to the NAS Rule, namely the Hierarchy Principle for industries and the Rank Size Rule for cities.

These stability results are even more remarkable given the substantial shuffling of both industries and city sizes between these two years. For the NAS Rule in particular, the results in equations 6 and 7 show that in spite of dramatic changes in both the cb-choice cities for specific industries and even the sizes of these cities themselves, the elasticity of substitution between the number and average size of these cb-choice cities across industries has remained essentially constant. While the corresponding log-linear relationship for the Rank Size Rule in equation 11 is not as sharp, the overall elasticity of substitution between city sizes and ranks in equations 12 and 13 has also remained essentially constant (and slightly larger in magnitude than for the classical model). Thus, while the underlying adjustment processes that preserve these relations remain to be established, it would appear that such processes must be relatively fast in comparison to this twenty-year span.

In addition, the joint stability of these three relations serves to reinforce the close relationships between them. With respect to the Rank Size Rule in particular, these relations suggest that rather than considering simple independent growth models of cities, such as Gibrat's Law and its extensions (see Gabaix and Ioannides 2004), better explanations of the skewed distribution of city sizes might be given in terms of the colocation behavior of populations and industries over time.

It should also be noted that while the present results for the NAS Rule involve only two points in time for a single country (Japan), this regularity appears to be far more robust. As mentioned at the beginning, the results of Hsu (2008) suggest that this same regularity can be seen in the United States as well. More generally, it appears that the NAS Rule is also evident in relatively self-contained subregions of nations. In particular, if one takes patterns of interregional travel behavior to define relatively self-sustained subsystems within nations (in the same way that commuting patterns have been used to define cities), then for Japan there is a natural nesting of four monopolar regional systems identified by their central cities as "Tokyo" $\supset$ "Osaka" $\supset$ "Nagoya," and "Tokyo" $\supset$ "Sapporo".[35] Our preliminary investigations show that the NAS Rule holds with roughly the same slope coefficient for the "Tokyo," "Osaka,"

35. Data on interprefectural passenger trips using mass transport modes were obtained from the Ministry of Land, Infrastructure, Transport and Tourism of Japan (2000).

"Nagoya," and "Sapporo" regions. These initial findings suggest that international comparisons of such regularities would perhaps be most meaningful by identifying self-contained economically comparable subregions for testing purposes. Such questions will be pursued further in subsequent research.

Finally, it should be emphasized that while the NAS Rule implies a regularity between the number and average size of cb-choice cities for each industry, it says little about the actual distribution of industries across cities. Hence from a regional policy perspective, neither the existence nor the stability of NAS by itself allows specific conclusions to be drawn about this distribution. However, when taken together with the closely related Hierarchy Principle, there are some policy implications that can be drawn. Indeed, if the Hierarchy Principle were to hold exactly, then the set of cb-choice cities for each industry i would be completely determined by the number of such cities, that is, would consist of the $\#U_i$ cities with largest industrial diversities. In such a case, cities k could only hope to attract new industries i for which k would then qualify as a cb-choice city for i and all cities with larger industrial diversity were already cb-choice cities. While such rigid rules are of course unrealistic, they nonetheless suggest that in regional systems where these regularities are sufficiently strong, cities are more likely to attract industries for which this addition would either create or enhance a meaningful local clustering of that industry and would be consistent with the current locational hierarchy for that industry. In particular, this suggests that smaller cities may be more likely to grow by attracting lower-order industries that would not be too isolated in that city. Such policy implications will be considered more fully in subsequent work.

Appendix

Industry Aggregation

There are 152 and 164 classifications in the three-digit manufacturing industries in 1981 and 2001 defined in the Establishments and Enterprise Census of Japan (1981, 2001). Industrial classifications have been basically disaggregated over the twenty-year period. Thus, in the present paper, the classifications in 2001 have been basically aggregated to those in 1981. Besides the conversions of the classifications between the two periods specified in the census, however, the following conventions have been adopted to make classifications at these two time points comparable.

1. Since "forged and cast steel manufacturing" (JSIC316) and "cast iron product manufacturing" (JSIC317) in 1981 have been aggregated to "ferrous metal machine parts and tooling products" (JSIC266) in 2001, we redefined JSIC316 to represent a union of JSIC316 and JSIC317 in 1981 and JSIC266 in 2001.

2. Since the union of "headgear manufacturing" (JSIC213) and "other apparel and textile accessory manufacturing" (JSIC215) in 1981 is equivalent to the union of "Japanese style apparel and socks ('tabi')" (JSIC155) and "other textile apparel and accessories" (JSIC156) in 2001, these are labeled JSIC215.

3. Since "wooden footwear manufacturing" (JSIC224) and "other wooden product manufacturing" (JSIC229) in 1981 have been aggregated to "miscellaneous manufacture of wood products, including bamboo and rattan" (JSIC169) in 2001, we let JSIC229 represent this aggregated classification.

Comments

Yannis M. Ioannides: This paper reconsiders the so-called Number-Average Size (NAS) Rule by investigating the rule by means of new definitions for cities and the presence of industries in cities. It is built around an empirical finding with data from Japan that shows that a regression of the average population size of cities, where an industry is present, against the number of such cities yields a very precisely estimated log-linear relationship. This paper differs from previous work by Mori, Nishikimi, and Smith (2008), which relies on a traditional definition of a city and of the presence of an industry.

To appreciate this difference, consider that urbanization and industrial concentration are often dramatized by means of pictures of countries (or land masses) from space. Clustering of lights and their brightness is very suggestive of clustering of economic activity. What conclusions can one draw about firms' location decisions from such pictures, or from the underlying geocoded data, when industrial concentration may not observe jurisdictional boundaries?

In the remainder of this comment I first outline the paper and discuss its key findings, then seek to put them in the perspective of the industrial agglomeration literature. I conclude with a critique of the findings and suggestions for future research.

Mori and Smith's paper in this volume reaffirms the validity of the NAS Rule, which they report first in Mori, Nishikimi, and Smith (2008). Mori and Smith use "clustered-based" choice cities of industries for two cross sections, 1980/1981 and 2000/2001, of Japanese data. Mori, Nishikimi, and Smith look only at choice-based cities of each industry, which are those hosting "significant" agglomeration of an industry. Mori and Smith use a different definition of Urban Employment Areas (UEAs) that expands the data to include smaller urban concentrations. That is, unlike Mori, Nishikimi, and Smith who use only Metropolitan Employment Areas (MEAs), which are UEAs with central municipality populations of at least 50,000, Mori and Smith include those with central municipality populations of at least 10,000. This broader definition yields 309 cities (UEAs) in 1980 and 258 cities in 2000 (compared with the smaller sets of 105 and 113 MEAs, respectively, used by Mori, Nishikimi, and Smith). Mori and Smith determine agglomerations of each industry that are significant by means of a statistical clustering technique, which they develop in two other papers (Mori and Smith 2009a, 2009b).

Specifically, the measures that are typically used for studies of industrial agglomeration are based on data collected according to jurisdictional boundaries. For the United States, these include states or local jurisdictions. The U.S. Bureau of the Census provides data at different spatial scales, such as U.S. metropolitan areas, U.S. census regions, and so on. Similar procedures are used elsewhere, which in the European Union involve NUTS (Nomenclature of Territorial Units for Statistics) and are designed to handle various spatial scales.

Mori and Smith use cluster analysis to detect whether the observed spatial distribution of establishments is not random and may be explained best by using statistical model selection criteria for finding the "best cluster scheme." For each of a number of statistical techniques, such as likelihood-ratio tests, Akaike's information criterion, and Schwartz's Bayesian Information Criterion, the authors compute the difference between the particular measure when it is based on the observed distribution and when it is based on complete spatial randomness. The best cluster scheme is the one that maximizes this difference.

This sounds straightforward, but there is a difficulty: the method requires defining spatial partitions, but the number of possible partitions of the space can be enormous. For this reason, these authors propose a cluster detection procedure that detects, in general, one cluster per industry. This designates a partition of the national economic space where an industry's concentration is significantly most pronounced.

I wish to explain this process a bit further. I adopt the authors' notation, according to which an economy's continuous (location) space Ω, is subdivided into disjoint municipalities, Ω_r, Ω_r, $\subseteq \Omega$ with municipalities indexed by the set $R = \{1, ..., k_R\}$. The municipalities partition the economy's space: $\cup_{r=1}^{k_R} \Omega_r = \Omega$. Suppose that establishment locations over space Ω may be described in terms of an industry-specific probability distribution function. Location decisions of different establishments in an industry may be treated as independent random samples from this unknown distribution. The class of all possible location models corresponds to the set of probability measures on Ω.

Suppose next that we identify groups of municipalities within which an industry's locational activity is more intense, that is possibly disjoint *clusters* of municipalities, defined by subsets of the index set $\boldsymbol{R}$: $C_j \subset R$, $j \in C = \{1,\ldots,$

$k_c\}$. All clusters make up a cluster scheme, **C**, which is a partition of the index set of municipalities, $\mathbf{C} = (R_0, C_{1,\ldots,} C_{k_C})$.

For example, let the jurisdictions be U.S. states. In that case, a cluster could be the group of the New England states {Massachusetts, Connecticut, Rhode Island, New Hampshire, Vermont, Maine}. The areal extent of cluster *Cj* is the union of the areas of all of its constituent jurisdictions, VCj = $\cup_{r\in C}^{k_R}$, the entire land area of New England. The probability that an establishment locates in New England is $p_C(j) = P_C(\Omega_{C_i}), j \in C$.

Mori and Smith define choice cities of industries and choice industries of cities as follows. In the institutional context of Japan, consider Urban Economic Areas, $\text{UEA}_k \subseteq U$ and cluster schemes for different industries. Let UEA_k overlap territorially with cluster C_j in cluster scheme **C**, $\text{UEA}_k \cap C_j \neq \emptyset$. Then, city UEA_k is said to be a choice city of industry *j*, that is, industry *j* establishments may be found in city *k*. Let U_j be the set of choice cities of industry *j*. Conversely, industry *j* is a choice industry of city *k*. Let the set of choice industries of city *k* be I_k. The more spatially diverse an industry is, the larger the number of its choice cities U_j (the location diversity of industry *j)*, and the less localized it is.

The authors construct cluster schemes for each of the 147 manufacturing industries for 1981 and 2001. Of these, for eight industries, complete spatial randomness could not be rejected. Of the remaining 139 industries, the percentages of establishments included in clusters range from 39.1 percent to 100 percent (with an average of 94.1 percent) in 2001, while the corresponding percentages in 1981 range from 51.8 percent to 100 percent (with an average of 95.7 percent).

Consider the ratio, for each industry, of the mean of employment among all cluster-based cities (the cities that are included in the industry's cluster, that is, the choice cities of that industry) to the mean of employment among those that are not in its cluster (the nonchoice cities of that industry, where the industry might be merely present) but do contain positive employment by plants of that industry. A similar measure can be defined in terms of establishment counts, instead of employment. Both these measures, when plotted are remarkably regular; see figures 2 and 3 in Mori and Smith. Both figures suggest that these magnitudes are skewed to the right. This regularity is tantalizing. It would behoove the authors to see whether these frequency distributions could be predicted on the basis of the sampling process that generates the respective measures.

Mori and Smith also show that, over a twenty-year period, industrial location decisions are subject to a lot of churning. This is illustrated by figures 4 to 7 in the paper. That is, during 1980/1981–2000/2001, smaller and less diversified cities tend to stay smaller and less diversified, and larger and more diversified cities also tend to stay larger and more diversified cities. At the same time, the locations at which agglomeration of industries take place vary quite a lot. Remarkably then, Mori and Smith find a statistical regularity—that is, a linear regression, across the three-digit manufacturing industries in their data set, of the log of the average size of industry choice cities against the log of their number gives a nearly perfect fit. This is the NAS Rule, pictured in figure 9, along with the outliers. They also find it to be remarkably stable over time. This stunning fit, which as the authors state should best be regarded as "curve fitting," rather than a genuine statistical model, is indistinguishable from perfect linearity. Furthermore, Holmes and Hsu (2009) report estimates of the coefficient of the average size using U.S. data (3-digit NAICS identifier for MSA and CMSAs in 2000) and the overall fit, at 0.7477 (standard deviation = 0.00255) and $R^2 = 0.9991$, respectively, that are remarkably close to those of Mori and Smith.[1]

The authors also report regressions along the lines of the Rank Size Rule. Figure 10 plots the logarithm of population size against the logarithm of the rank, for the same definition of choice-based cities that was used in their NAS Rule, along with the corresponding linear regression lines. The resulting fits for the Rank Size Rule are very different, in my opinion, from those of the NAS Rule. As the authors state, the overall slope of each regression is close to –1.0, but they are significantly different, as I see it, from the classical Rank Size Rule. As they put it, "but it is clear for the regression line shown in plot e of figure 10 for 2000/2001 data that this slope is in fact a compromise between the slopes" for the middle range and the two extreme ones. The regression line cuts through what is roughly a concave plot.

Here we have a genuine instance of "glass half full" versus "glass half empty." The authors state that "while no definitive conclusions can be drawn from such limited observations," they do suggest that the theoretical relations between the NAS Rule and both the Rank Size Rule and Hierarchy Principle developed in Mori, Nishikimi, and Smith (2008) are empirically most evident for cities in this dominant central range." In comparison, Duranton (2007) and Rossi-Hansberg and Wright (2007) propose good theoretical explanations for the concavity of the city size distribution.

Specifically, Mori, Nishikimi, and Smith offer an elegant result on equivalence between NAS and Rank Size Rules (see Mori, Nishikimi, and Smith 2008, Theorem 1). A critical step that allows them to prove their Theorem 1 is that

1. NAICS = North American Industry Classification System; MSA= metropolitan statistical area; CMSAs = consolidated metropolitan statistical areas; S.D. = standard deviation.

when cities are indexed, the indexes of cities where an industry is present (its choice cities) form an interval. Specifically, let industry types be defined by index $i \in \mathbf{I}$, with industry i occupying a measurable subset $\mathbf{R}_i \subset \mathbf{R}$ of the cities in $\mathbf{R}$. Let $\mathbf{R}_i$ be an interval, $\mathbf{R}_i = [0, r_i]$, where r_i denotes the rank of the smallest population units occupied by industry i. This assumption encompasses a strict version of the Hierarchy Principle. In general, this principle holds that if an industry is present in a city, it would also be present in all cities that are larger than itself. Intuitively, there are cities with gas stations only, but a larger city with an opera will have gas stations, as well.

Let n_i denote the number of cities where industry i is present, and the number of choice cities of industry i as

$$n_i = \int_{R_i} dr.$$

The average size of choice cities of industry i is

$$\bar{R}_i = (1/n_i)\int_{R_i} \rho(x)\, dx.$$

Cities are ranked by their sizes: rank(r) = $\rho(r)$. The empirical NAS Rule is expressed as

$$\text{If } \bar{R}_i = \alpha n_i^{-\beta}, \text{ then } \rho(r) = \alpha r^{-\beta} \Leftrightarrow \bar{R}_i = \alpha(1-\beta)^{-1} n_i^{-\beta}.$$

Taking logs of both sides in the last equation above gives the log of $\bar{R}_i$ as a linear function of the log of n_i. Theorem 1 states that for any economy for which Christaller's Hierarchy Principle holds, the Rank Size Rule and the NAS Rule are equivalent.

In contrast, Rossi-Hansberg and Wright (2007) use a complete theory of urban structure and growth to argue that larger cities likely operate in industries that experienced a history of above-average productivity shocks, and thus they can be expected to grow slower than average in the future, while the opposite is true of smaller cities. Furthermore, urban growth rates exhibit reversion to the mean, which implies "that the log rank-log size relationship will in general (apart from particular realizations of the shocks) be concave or, in other words, that the invariant distribution for city sizes has thinner tails than a Pareto distribution with coefficient 1."[2]

The authors' linking the NAS Rule with the Rank Size Rule does not strengthen their finding, in my opinion, for two reasons. One is that the Rank

2. Rossi-Hansberg and Wright (2007, p. 612).

Size Rule might be in the eye of the beholder. A second is that the Rank Size Rule is not an immutable fact. Ioannides and others (2008) use international city size data to show that adoption of information and communication technologies causes, ceteris paribus, an increased concentration of the city size distribution and thus a decreased Zipf's coefficient. Still, this misgiving of mine should not be held against the significance of the authors' contribution that is reported in the paper.

Still, the conceptual interrelationships between the NAS and Rank Size Rules, on the one hand, and the Hierarchy Principle, on the other, are very interesting and too tantalizing to ignore. To model them, one could start from a model of plant location, perhaps along the lines of Ellison and Glaeser (1997). Such a model assigns probabilities to location decisions of different firms, from which probabilities can be computed for different realizations of strings of zeroes and ones. Such strings feature prominently in the authors' graphical representations of the Hierarchy Principle. Location choices that depend on industry presence and city size could provide a link from the Hierarchy Principle, a qualitative relationship, to the NAS and Rank Size Rules. In this connection, theoretical results by Hsu (2008), however special, are very promising in that they can provide an overarching theme linking qualitative and quantitative aspects of location.

I conclude that I am most amazed by and respectful of the authors' NAS Rule, both as reported here as well as in Mori, Nishikimi, and Smith (2008). Furthermore, the consistency of the findings across different definitions of urban areas, not to mention the care with which the search of cluster schemes is implemented, is also impressive. I am all the more anxious and hopeful to see that a full theory underlying these findings would be developed. The potential for informing the design for urban and regional development policies is also considerable. Policies must respect the fact that industries tend to cluster and in ways that are related to the size of the urban areas hosting them.

John M. Quigley: The careful and well-documented paper by Tomoya Mori and Tony Smith makes four distinct and interesting contributions. First, it presents and explicates a new measure of the regularity of city types within a country, the so-called Number-Average-Size Rule. Second, it shows how this new measure is related to a group of standard and well-known economic and geographic measures of empirical regularity across cities. Third, the paper demonstrates that the Number-Average-Size measure is generated by agglomerative factors in the economic geography of regions, not simply by chance.

More specifically, it shows that the empirical regularities observed in Japan at a couple of points in time do not arise from some competing "dartboard model" or some random set of fluctuations. Fourth, it presents a comprehensive description of the computation of these measures, and the hierarchical significance of these measures, for all the 3,200 municipalities in Japan.

What is the new measure of the distribution of city sizes, the Number-Average-Size Rule? The rule asserts that the logarithm of city size (that is, population) is a linear function of the logarithm of the number of industries that are represented in that city. The threshold for representation is pretty clear, and the statistical model employed by the authors fits the data remarkably well, at least for Japanese cities. Essentially all the variation in Japanese city sizes is explained by the heterogeneity of the industrial composition of those cities. This is true for two cross sections of municipalities measured two decades apart, in 1980 and in 2001. Moreover, the estimated slopes of the regression relationships are virtually identical in these cross sections measured twenty years apart.

How is this new measure related to standard measures of the urban hierarchy? Seventy-five years ago, Walter Christaller developed a simple principle of urban hierarchies that was based upon introspection and his close observation of villages, towns, and cities in Germany. In villages, the grain grown in the surrounding farms was auctioned in local markets. In the next tier of places, towns, the grain was milled into flour and shipped onto larger towns. But some was baked into bread and sold in those small towns and also in the villages that supplied the grain in the first place. From these observations came the specific hunch that all economic activities in a city will also be found in cities of larger sizes.

And from this, Christaller's famous Rank Size Rule follows. The Rank Size Rule holds that the logarithm of city size is linearly related to the logarithm of the rank of that city in the hierarchy of cities. Christaller formulated this law on the basis of introspection and the observations of rural life. From purely an empirical viewpoint, the Rank Size Rule has also been remarkably durable. It has been applied to the urban hierarchy of many countries in many time periods (see Berliant 2008 for a review). And there has been great attention paid to theoretical constructs that would justify the remarkable regularities observed. (See Simon 1955; Gabaix 1999; Gabaix and Ioannides 2004.)

Specifically, Mori and Smith demonstrate in their paper that a specific weak formulation of Christaller's hunch unifies the Number-Average Size Rule and the Rank Size Rule. Specifically, assume that those industries in cities of a given industrial diversity will also be found in cities of higher industrial diversity.

Under this assumption, the authors show that Christaller's Rank Size Rule is isomorphic to their Number-Average Size Rule.

More important, the assumption that a large city like Tokyo has *A* slots for different industries while a smaller city like Osaka has *B* slots for different industries ($A > B$) facilitates a direct test of the extent to which the joint frequency distribution of population in cities and the number of industries in those cities could arise by chance.

The authors take great care to standardize the linkage between population and industries for Japanese cities—standardization is by characteristics of transport systems and demographics. The empirical analysis demonstrates rather convincingly that the Number-Average Size Rule does not arise from the random or quasi-random location of industries or from quirks in geography or the transport system.

The authors demonstrate that there have been enormous changes in the hierarchy of Japanese cities. Below the thirty largest cities, there has been a great deal of movement—with some cities going up the hierarchy and some moving downward over relatively short periods of time. Despite this vast churning of cities in the hierarchy, the Number-Average Size Rule operates in any cross section. This is the most compelling finding in the paper.

My principal criticism of the analysis so far is that it is essentially mechanical. So far, the authors have not been able to dig into the economic or social behavior of households and firms that might give rise to these spatial regularities.

If it is true that city size is associated with a more heterogeneous industrial structure, is economic output similarly associated? Is output per worker, that is, productivity, associated as well? Can better city organization be inferred from a city's deviation from its predicted placement according to the Number-Average Size Rule? There is a hint of an argument about this in the paper, but it is not taken very far.

But we could well imagine taking the Number-Average Size Rule seriously in economic research, not just in spatial or geographical descriptions. Recall, there are a number of studies that have explored the variations in the Rank Size Rule across countries—to what extent does the exponential decay parameter vary across the hierarchy in different countries? With policies? With the political economy? With favoritism accorded to the capital city?

One could imagine that the parameter of the Number-Average Size Rule, the exponential decay, varies across countries—with the level of development or the extent of active regional policies to attract (or direct) types of industries to different places. To what extent is this measure of industrial structure consistent with agglomerative urbanization economies? In Japan or in comparison with other places?

At this point, the challenge to the authors is to test whether the careful spatial measurement in this paper is useful in understanding the behavior of the economic and social agents in the urban hierarchy.

References

Akaike, H. 1973. "Information Theory as an Extension of the Maximum Likelihood Principle." In *Second International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki, pp. 267–81. Budapest: Akademiai Kiado.

Berliant, Marcus. 2008. "Central Place Theory." In *The New Palgrave Dictionary of Economics*, 2nd ed. New York: Palgrave Macmillan.

Besag, J., and J. Newell. 1991. "The Detection of Clusters in Rare Diseases." *Journal of the Royal Statistical Society*, *Series A* 154: 143–55.

Black, D., and V. Henderson. 2003. "Urban Evolution in the USA." *Journal of Economic Geography* 3: 343–72.

Christaller, W. 1966. *Central Places in Southern Germany*, translated by C. W. Baskin. Upper Saddle River, N.J.: Prentice Hall. Originally published as *Die zentralen Orte in Süddeutschland* (Gustav Fischer, 1933).

Dasgupta, A., and A. E. Raftery. 1998. "Detecting Features in Spatial Point Processes with Clutter via Model-Based Clustering." *Journal of the American Statistical Association* 93: 294–302.

Duranton, G. 2007. "Urban Evolutions: The Fast, the Slow, and the Still." *American Economic Review* 97 (1): 197–221.

Eaton, J., and Z. Eckstein. 1997. "Cities and Growth: Theory and Evidence from France and Japan." *Regional Science and Urban Economics* 27: 443–74.

Ellison, Glenn, and Edward L. Glaeser. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy* 105 (5): 889–927.

Gabaix, X. 1999. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics* 114: 737–67.

Gabaix, X., and Y. M. Ioannides. 2004. "The Evolution of City Size Distributions." In *Handbook of Regional and Urban Economics*, vol. 4, *Cities and Geography*, edited by J. V. Henderson and J.-F. Thisse, pp. 2341–378. Amsterdam: North-Holland.

Gangnon, R. E. and M. K. Clayton. 2000. "Bayesian Detections and Modeling of Spatial Disease Clustering." *Biometrics* 56: 922–35.

———. 2004. "Likelihood-Based Tests for Localized Spatial Clustering of Disease." *Environmetrics* 15: 797–810.

Holmes, Thomas J., and Wen-Tai Hsu. 2009. "Optimal City Hierarchy: A Dynamic Programming Approach to Central Place Theory." Working Paper. Chinese University of Hong Kong, Department of Economics (January).

Hsu, Wen-Tai. 2008. "Central Place Theory and Zipf's Law." Working Paper. Chinese University of Hong Kong (December).

Ioannides, Yannis M., Henry G. Overman, Esteban Rossi-Hansberg, and Kurt Schmidheiny. 2008. "The Effect of Information and Communication Technologies on Urban Structure." *Economic Policy* 54: 203–42.

Japan Statistics Bureau. 1980. *Population Census*.

———. 1981. *Establishments and Enterprise Census of Japan*.

———. 2000. *Population Census*.

———. 2001. *Establishments and Enterprise Census of Japan*.

Kanemoto, Y., and K. Tokuoka. 2002. "The Proposal for the Standard Definition of the Metropolitan Area in Japan." [In Japanese.] *Journal of Applied Regional Science* 7: 1–15.

Kontkanen, P., and P. Myllymäki. 2005. "Analyzing the Stochastic Complexity via Tree Polynomials." Technical Report 2005-4. Helsinki Institute for Information Technology.

Kulldorff, M. 1997. "A Spatial Scan Statistic." *Communications in Statistics: Theory and Methods* 26: 1481–496.

Kulldorff, M., and N. Nagarwalla. 1995. "Spatial Disease Clusters: Detection and Inference." *Statistics in Medicine* 14: 799–810.

Ministry of Land, Infrastructure, Transport and Tourism of Japan. 2000. *Inquiry for Passenger Trip Flow among Prefectures*.

Mori, Tomoya, Koji Nishikimi, and Tony E. Smith. 2008. "The Number-Average Size Rule: A New Empirical Relationship between Industrial Location and City Size." *Journal of Regional Science* 48 (1): 165–211.

Mori, Tomoya, and Tony E. Smith. 2009a. "A Probabilistic Modeling Approach to the Detection of Industrial Agglomerations." Working Paper.

———. 2009b. "Spatial Coordination of Population and Industrial Agglomerations: The Central Place Theory and the City Size Regularities Revisited." Working Paper.

National Statistics Center of Japan. 2009. "Portal Site of Official Statistics of Japan." Statistics Bureau, Ministry of Internal Affairs and Communication (www.e-stat.go.jp/SG1/estat/eStatTopPortalE.do).

Overman, H., and Y. M. Ioannides. 2001. "Cross-Sectional Evolution of the US City Size Distribution." *Journal of Urban Economics* 49: 543–66.

Rossi-Hansberg, Esteban, and Mark L. J. Wright. 2007. "Urban Structure and Growth." *Review of Economic Studies* 74: 597–624.

Schwarz, G. 1978. "Estimating the Dimension of a Model." *Annals of Statistics* 6: 461–64.

Simon, Herbert. 1955. "On a Class of Skew Distribution Functions." *Biometrika* 44: 425–40.

Smith, T. E., and T. Mori. 2009. "On the Model-Based Clustering for the Detection of Industrial Agglomerations." Working Paper.

United Nations. 2007. *World Urbanization Prospects*.

U.S. Office of Management and Budget. 2000. "Standards for Defining Metropolitan and Micropolitan Statistical Areas." *Federal Register* 65 (249) (December 27).