Urban Growth and the Simon Process: the Japanese Case

Toshitaka Fukiharu
(Faculty of Economics, Hiroshima University)

January, 2006

Introduction

There are two approaches in the urban economics: normative and positive ones. One of the main themes in Japan for the former approach is the transfer of the capital. This theme has been often discussed in Japan and South Korea since the population is overconcentrated on Tokyo and Seoul, while rural areas suffer from depopulation. Some argue for the transfer, pointing out the high living costs including land price and atmospheric pollution there. The others argue against it, pointing out that the high living cost reflects the high productivity including informational value there. In Japan, by political reasons, the merging of the depopulated towns are under way. The aim of this paper is not to discuss whether the city size distribution in a country is desirable or not. Rather, the theoretical explanation of the city size distribution is the main concern. Thus, the positive approach is adopted in this paper.

It is well-known that in any country all over the world the city size distribution exhibits the Pareto distribution with parameter value approximately one (Zipf [1949], Rosen and Resnick [1980], and Dobkins and Ioannnides [1996]). On the one hand, purely economic explanation has been attempted (Henderson [1988], Fujita, Krugman and Venables [1999]). On the other hand, somewhat more general (non-economic) explanation has also been attempted (Simon [1955], Ijiri and Simon [1977]). In this paper, the latter approach by Simon is adopted. In Section 1 of this paper, the comparison of the US city size distribution and the Japanese one is made, explaining the necessity to modify the Simon process. In Section 2, the Mathematica programming of traditional Simon process is explained, while other examples of Pareto distribution are constructed. In Section 3, the (modified) two-stage Simon process is proposed and a Mathematica command function is constructed for simulations. Utilizing this command function, the simulations on this process are conducted. The final section concludes this paper.
1. The Comparison of the US City Size Distribution and the Japanese One

1.1 The US City Size Distribution

Fujita, Krugman and Venables [1999] applied one of the purely urban-economic analysis to the explanation of the US city size distribution. According to their conclusion, the result is not satisfactory. The US city size distribution can be easily constructed utilizing Statistical Abstracts of the United States 2002. If the cities with more than 100,000 population are selected and the set of their populations, city sizes, is denoted as "dataUS". The data contains 239 cities, where the largest city size is 8,008,000 (New York City) and the smallest one is 100,000, while "dataUS" contains 153 different elements: city sizes.

\[
\{\text{Length}[\text{dataUS}], \text{Max}[\text{dataUS}], \text{Min}[\text{dataUS}], \text{dataUS}2 = \text{Union}[\text{dataUS}]; \text{Length}[\text{dataUS2}]\}\n\]

\[\{239, 8008, 100, 153\}\]

The following set, "dataUS2A", which is the set of pairs, \(\{S,N\}\), where \(S\) is the city size in "dataUS", and \(N\) is the number of city sizes in "dataUS" at least equal to \(S\).

\[
\text{dataUS2A} = \text{Table}[\{\text{dataUS2}[[k]], \text{Length}[\text{Select}[\text{dataUS}, \# \geq \text{dataUS2}[[k]] &]]\}], \{k, 1, \text{Length}[\text{dataUS2}]\}\]

The following set, "dataUS3A", is the logarithmic transformation of dataUS2A, and it is graphically depicted as in what follows: figA1.

\[
\text{dataUS3A} = \text{Log}[\text{dataUS2A}]; \text{figA1} = \text{ListPlot}[\text{dataUS3A}, \text{AxesLabel} \rightarrow \{\text{"log } S\}, \text{"log } N\}, \text{PlotJoined} \rightarrow \text{True}, \text{PlotLabel} \rightarrow \text{"Figure City(US)"}];
\]

Utilizing the regression analysis through the least square method, we have the following functional relation:

\[
\log N = 11.6203 - 1.33573 \log S \tag{1}
\]

Note that \(R^2\) value in this regression is quite high: \(R^2 = 0.994036\).
Zipf[1949] found that for the US city size distribution

\[ N = kS^{-\alpha} \]  \hspace{1cm} (2)

holds where \( \alpha \) is approximately 1. The distribution of \( \{S, N\} \) which satisfies (2) is called the Pareto distribution.

### 1.2 The Japanese City Size Distribution

Rosen and Resnick [1980] found that for any country all over the world the city size distribution exhibits the Pareto distribution with parameter \( \alpha \) approximately one. In this subsection, it is examined if the same result obtains for the Japanese case. According to *Japan Statistical Yearbook 2004*, there are 671 cities in Japan. Among the collection of Japanese city sizes, "data J2", the largest one is 8,130,408 (Tokyo) and the smallest one is 5,941 (Utashiuchi).

Following the same procedure as in subsection 1.1, the set, "data J2A", the set of pairs, \( \{S, N\} \), where \( S \) is the city size in "data J2", and \( N \) is the number of city sizes in "data J2" at least equal to \( S \).

Following the same procedure as in subsection 1.1, the set, "data J2A", the set of pairs, \( \{S, N\} \), where \( S \) is the city size in "data J2", and \( N \) is the number of city sizes in "data J2" at least equal to \( S \).

The following set, "data J3A", is the logarithmic transformation of "data J2A", and it is graphically depicted as the thick solid curve in what follows: fig J1.

```mathematica
Regress[dataUS3A, \{1, x\}, x]
\begin{array}{lrcrr}
\text{ParameterTable} & 1 & 11.6203 & 0.0474219 & 245.041 & 0. \\
\text{x} & -1.33573 & 0.00841937 & -158.65 & 7.30135 \times 10^{-170}.
\end{array}
\text{RSquared} \rightarrow 0.994036, \text{AdjustedRSquared} \rightarrow 0.993997, \text{EstimatedVariance} \rightarrow 0.00646447,
\begin{array}{lrcrr}
\text{ANOVATable} & \text{Model} & 1 & 162.709 & 162.709 & 25169.7 & 0. \\
\text{Error} & 151 & 0.976135 & 0.00646447 & \\
\text{Total} & 152 & 163.685 & \\
\end{array}
\text{Zipf}[1949] \text{found that for the US city size distribution}
\begin{equation}
N = kS^{-\alpha}
\end{equation}
\text{holds where} \alpha \text{is approximately 1. The distribution of} \{S, N\} \text{which satisfies (2) is called the Pareto distribution.}

\text{1.2 The Japanese City Size Distribution}

Rosen and Resnick [1980] found that for any country all over the world the city size distribution exhibits the Pareto distribution with parameter \( \alpha \) approximately one. In this subsection, it is examined if the same result obtains for the Japanese case. According to *Japan Statistical Yearbook 2004*, there are 671 cities in Japan. Among the collection of Japanese city sizes, "data J2", the largest one is 8,130,408 (Tokyo) and the smallest one is 5,941 (Utashiuchi).

Following the same procedure as in subsection 1.1, the set, "data J2A", the set of pairs, \( \{S, N\} \), where \( S \) is the city size in "data J2", and \( N \) is the number of city sizes in "data J2" at least equal to \( S \).

Following the same procedure as in subsection 1.1, the set, "data J2A", the set of pairs, \( \{S, N\} \), where \( S \) is the city size in "data J2", and \( N \) is the number of city sizes in "data J2" at least equal to \( S \).

The following set, "data J3A", is the logarithmic transformation of "data J2A", and it is graphically depicted as the thick solid curve in what follows: fig J1.

```mathematica
N = k S^{-\alpha}
```
Utilizing the regression analysis through the least square method, we have the following functional relation:

\[ \log N = 17.2024 - 1.03518 \log S \tag{3} \]

Note that \( R^2 \) value in this regression is high: \( R^2 = 0.95486 \).

Regress[\( \text{dataJ3A}, \{1, x\}, x \)]:

\[
\begin{array}{cccc}
\text{ParameterTable} & \text{Estimate} & \text{SE} & \text{TStat} & \text{PValue} \\
1 & 17.2024 & 0.0985756 & 174.509 & 4.6265940512 \times 10^{-560} \\
x & -1.03518 & 0.00870195 & -118.96 & 2.826528167349 \times 10^{-452} \\
\end{array}
\]

\( \text{RSquared} \rightarrow 0.95486, \text{AdjustedRSquared} \rightarrow 0.954792, \text{EstimatedVariance} \rightarrow 0.0434652, \text{ANOVA} \rightarrow \)

\[
\begin{array}{cccc}
\text{DF} & \text{SumOfSq} & \text{MeanSq} & \text{FRatio} & \text{PValue} \\
\text{Model} & 1 & 615.095 & 615.095 & 2.826528167349 \times 10^{-452} \\
\text{Error} & 669 & 29.0782 & 0.0434652 & \\
\text{Total} & 670 & 644.173 & \\
\end{array}
\]

Between the two results, there is a small difference. The degree of statistical fit for the linear relation in (1) is better than the one in (3). From figJ1, it appears that there is non-linear functional relation between \( \log S \) and \( \log N \) for Japanese case, so that the non-linear least square method is applied to "dataJ3A".

\[
\begin{align*}
e1 &= \text{Fit[\( \text{dataJ3A}, \{1, x\}, x \)}, \\
\text{fige1} &= \text{Plot[e1, \{x, 8, 16\}, \text{PlotStyle} \rightarrow \text{Dashing[\{0.01, 0.01\}]}, \text{DisplayFunction} \rightarrow \text{Identity}];} \\
f1 &= \text{Fit[\( \text{dataJ3A}, \{1, x, x^2\}, x \)}; \\
\text{figf1} &= \text{Plot[f1, \{x, 8, 16\}, \text{DisplayFunction} \rightarrow \text{Identity}];} \\
\text{figJ2} &= \text{Show[fige1, figf1, figJ1, \text{DisplayFunction} \rightarrow \$DisplayFunction,} \\
& \text{AxesLabel} \rightarrow \{"log S", "log N"\}, \text{PlotLabel} \rightarrow \text{"Figure City(J1AA)"};}
\end{align*}
\]

In figJ2, the dashed line is expressed by (3) and the \emph{thin} solid curve depicts the following equation:

\[ \log N = -0.626421 + 2.02444 \log S - 0.130256 \log S^2 \tag{4} \]

where \( R^2 \) is 0.988933.
Utilizing the regression analysis through the least square method, we have the following relation:

\[
\log N = 20.6306 - 1.31383 \log S
\]

Note that the \( R^2 \) value in this regression is high: \( R^2 = 0.988074 \). It may be said that the linearity holds on the functional relation between \( \log S \) and \( \log N \) when only the larger-sized cities are selected.

Moving back to the US case, the extended US city size distribution can be easily constructed utilizing *Statistical Abstracts of the United States 2002*, by adding cities with smaller population. If the cities with more than 10,000 population are selected and the set of their populations, city sizes, is denoted as "dataUS1", the set contains 2642 cities.
Utilizing the regression analysis through the least square method, we have the following functional relation:

\[
\log N = 18.5346 - 1.13681 \log S
\]

Note that the $R^2$ value is high in this regression: $R^2 = 0.98608$.

From the comparison between Figure City(US) and Figure City(A1) and the one between Figure City(Japan) and Figure City(J1) the following tendency may be observed.

- [O1] For both countries, as the data set is enlarged by containing smaller-sized cities in the data set, the absolute value of estimated parameter for the Pareto distribution becomes smaller.
- [O2] For both countries, as the data set is enlarged by containing smaller-sized cities in the data set, the non-linearity of the functional relation between $\log S$ and $\log N$ is stronger, while the linearity holds when only the larger-sized cities are selected.
- [O3] The degree of concavity for the functional relation between $\log S$ and $\log N$ is stronger in Japan than in the US.
2. Simon Process

As pointed out in Fujita, Krugman and Venables [1999, Chapter 12], Simon Process appears to create the linearity between log $S$ and log $N$, where the Simon Process is the growing process of cities. This theoretical result, however, is not necessarily compatible with the actual city distribution. As noted in the previous section, the actual city distribution does not necessarily exhibit the linearity of functional relation between log $S$ and log $N$ over the whole region, but only on the subset of the whole region. To be precise, it exhibits the non-linearity with the estimated parameter of Pareto distribution almost equal to 1. Second, when larger-sized cities are selected it exhibits the linearity, while the estimated parameter is larger than 1. This observation of actual city distribution prompts us to modify the Simon Process. In this section, the Mathematica programming of traditional Simon process is explained, while other examples of Pareto distribution are constructed. The modification of the traditional Simon Process is attempted in section 3, incorporating the possibility of the break-up of the largest city in the growing mechanism of cities.

2.1 Programing of Traditional Simon Process

Simon [1955] constructed a general growing process of "lumps". In what follows, we take "city" as an example of "lump" and regard the dynamic Simon Process as the collection of city distributions from the initial period to the last period, T. In this process, as time proceed from current period to the next period, the probability of the creation of a new city (base unit, city size=1) in the next period is given by $\pi_1$, and the already existing cities grow in the next period with the probability proportionally to the population. As an example, consider the following special case.

Suppose that $\pi_1=0.1$, and currently, there are 8 base unit cities, 2 cities with city size 2, and one city with city size 4. The city size distribution is expressed as $\{(1,8),(2,2),(4,1)\}$. In this case, the probability of the creation of base unit city in the next period is 0.1, the probability of the expansion of existing base unit city to the city size 2 is $0.9 \times 1 \times 8 / (1 \times 8 + 2 \times 2 + 4 \times 1)$, the probability of the expansion of the existing city with city size 2 to the city size 3 is $0.9 \times 1 \times 8 / (1 \times 8 + 2 \times 2 + 4 \times 1)$, and the probability of the expansion of the existing city with city size 4 to the city size 5 is $0.9 \times 4 \times 1 / (1 \times 8 + 2 \times 2 + 4 \times 1)$. The program for this Simon Process is constructed as follows. The city size distribution, arbitrarily given in this period, $\{(1,8),(2,2),(4,1)\}$, is denoted as "list":

```mathematica
list = {{1, 8}, {2, 2}, {4, 1}};
```

We start with the expansion of the city size distribution to the following set, "data1":

```mathematica
{{0, 0}, {1, 8}, {2, 2}, {3, 0}, {4, 1}, {5, 0}}.
```

```mathematica
list1 = Table[list[[i, 1]], {i, 1, Length[list]}];
list2 = Table[i, {i, 0, Max[list1] + 1}];
list3 = Complement[list2, list1];
data1 = Union[Table[{list3[[i]], 0}, {i, 1, Length[list3]}], list]
```

```mathematica
{{0, 0}, {1, 8}, {2, 2}, {3, 0}, {4, 1}, {5, 0}}
```

Going back from "data1" to "list", it is renamed as "data1a".

```mathematica
data1a = Select[data1, #[[2]] > 0 &]
```

```mathematica
{{1, 8}, {2, 2}, {4, 1}}
```

The identifying function, $f[i]$, is defined, which shows the placing (position) in "data1" of the i-th element in "data1a". For example, the 2nd element of "data1a" corresponds to the 3rd element of "data1".

```mathematica
f[i_] := Position[data1, Intersection[{data1a[[i]]}, data1][[1]]][[1]]
```
The set of all the existing city sizes is defined by "scale": \{1,2,4\}.

\[
\text{scale} = \text{Table}[\text{data}\_1[[i, 1]], \{i, 1, \text{Length}\[\text{data}\_1]\}]
\]

\{1, 2, 4\}

The set all the numbers of cities corresponding to each city size is defined as "density": \{8,2,1\}.

\[
\text{density} = \text{Table}[\text{data}\_1[[i, 2]], \{i, 1, \text{Length}\[\text{data}\_1]\}]
\]

\{8, 2, 1\}

The set of all the products of each city size and the number of that city size is defined as "data2": \{8,4,4\}.

\[
\text{data2} = \text{Table}[\text{data}\_1[[i, 1]] \times \text{data}\_1[[i, 2]], \{i, 1, \text{Length}\[\text{data}\_1]\}]
\]

\{8, 4, 4\}

The set of all the sums of elements in "data2" up to the i-th element, i=1,2,3, is defined as "data3": \{8,12,16\}.

\[
\text{data3} = \text{Table}[\text{Sum}[\text{data2}[[i]], \{i, 1, k\}], \{k, 1, \text{Length}\[\text{data2}\]\}]
\]

\{8, 12, 16\}

Prepending 0 in "data3", adding 1 to the resulting new set, and finally deleting the last element of the resulting set, we obtain "data4": \{1,9,13\}, which is useful later.

\[
\text{data4} = \text{Drop}[\text{Prepend}[\text{data3}, 0] + 1, \{\text{Length}\[\text{data3}\] + 1\}]
\]

\{1, 9, 13\}

Supposing 1×8+2×2+4×1 as 'the total population', "population" is defined by the last element of "data3": 16.

\[
\text{population} = \text{Last}[\text{data3}]
\]

16

It was assumed in this specified case that the base-unit city is created with the probability \(p_1=0.1\), while one of the existing cities, say, the one with city size i, expands to the city with city size i+1, with the probability \(1-p_1=0.9\). In order to examine which case takes place, Mathematica functions, "Random" and "BernoulliDistribution", are utilized. Indeed, \text{Random}[\text{BernoulliDistribution}[p]] produces 1 with the probability p and 0 with the probability \(1-p\). Now, If it produces 0, we proceed to the selection of one element from \{1,2,...,16\}. If the element belongs to \{1,2,...,8\}, we assume that the city with city size 1 expands in the next period. If the element belongs to \{9,...,12\}, we assume that the city with city size 2 expands in the next period. If the element belongs to \{13,...,16\}, we assume that the city with city size 4 expands in the next period. In what follows, "data5" is the outcome for the command, while if \text{Random}[\text{BernoulliDistribution}[p]]=1, 0 is selected.

\[
\text{data5} = \text{If}[\text{Random}[\text{BernoulliDistribution}[0.9]] > 0, \text{Random}[\text{Integer}, \{1, \text{population}\}], 0]
\]

8

If "data5" is 0, all the elements in "data5"-"data4" are negative. If "data5" belongs to \{1,2,...,8\}, one of the elements in "data5"-"data4" is positive. If "data5" belongs to \{9,...,12\}, two of the elements in "data5"-"data4" are positive. Finally,
if "data5" belongs to \(\{13,...,16\}\), three of the elements in "data5"-"data4" are positive. Thus, "data6": the number of the positive elements, implies the position of the city, which expands in the next period.

\[
data6 = \text{Length}[\text{Union}[\text{Select}[\text{data5} - \text{data4}, \# \geq 0]]]
\]

1

If "data6"=0, we modify the second element (city size 1) of "data1" and add 1 to the number of the city. If "data6"=2, we modify the third element (city size 2) and the fourth element (city size 3) of "data1", and subtract 1 from the number of the city corresponding to the third element and add 1 to the number of the city corresponding to the fourth element. Thus, "data7", say \(\{2,3\}\), implies the set of positions of the consecutive two elements in "data1", which must be modified in this period.

\[
data7 = \text{If}[\text{data6} = 0, (1, 2), (\text{f[data6]}, \text{f[data6]} + 1)]
\]

\(\{2, 3\}\)

Indeed, from "data1" two consecutive elements are deleted by a Mathematica function, "Drop", to obtain "data8", say, \(\{\{0,0\},\{3,0\},\{4,1\},\{5,0\}\}\), for the modification of "data1".

\[
data8 = \text{Drop[\text{data1}, \text{data7}]}\]

\(\{\{0,0\},\{3,0\},\{4,1\},\{5,0\}\}\)

We proceed to the addition of 1 to the number of city, corresponding to the first element of "data7", and subtraction of 1 from the number of the city corresponding to the second element of "data7". Union of this modified "data7" and "data8" is "data9", say, \(\{\{0,0\},\{1,7\},\{2,3\},\{3,0\},\{4,1\},\{5,0\}\}\).

\[
data9 = \text{Union}[\text{data1}[[\text{data7}[[1]]]] - \{0,1\}, \text{data1}[[\text{data7}[[2]]]] + \{0,1\}], \text{data8}]
\]

\(\{\{0,0\},\{1,7\},\{2,3\},\{3,0\},\{4,1\},\{5,0\}\}\)

Selection of the elements, whose number of the city is positive, "data10", say, \(\{\{1,7\},\{2,3\},\{4,1\}\}\), is the city distribution in the next period.

\[
data10 = \text{Select[\text{data9}, \#[[2]] > 0 \&]}\]

\(\{\{1,7\},\{2,3\},\{4,1\}\}\)

This process is repeated from the 1st period to the last period \(T\). This is the traditional Simon Process. While the example was specified above, the accompanying Mathematica programming of the process is so general that the collection of the above series of programs is a Mathematica command function, which computes the next period's city size distribution, given \(\pi_1\) and the current period's city size distribution.

2.2 Other Examples of Pareto Distribution

Simon Process is a very general growing process of "lumps". Simon [1955] pointed out that a lot of distributions, such as the one of the numbers of words used in a book, or the one of the numbers of published papers by researchers in a technical journal, exhibit the Pareto distribution. He proved that if the distribution by Simon Process is convergent, then the limit distribution is a Pareto distribution. In this subsection, as an other example, the stock price distribution is examined. On January 4, 1988, there were 1202 stock prices in Tokyo Stock Exchange. The set of all the stock prices is denoted as "dataS1". The lowest stock price is 185 yen, the highest one is 2,080,000 yen.
\{\text{Length}[\text{dataS1}], \text{Min}[\text{dataS1}], \text{Max}[\text{dataS1}], \text{dataS2} = \text{Union}[\text{dataS1}]; \text{Length}[\text{dataS2}]\}

\{1202, 185, 2080000, 554\}

The following set, "dataS3", the set of pairs, \{S,N\}, where S is the stock price, and N is the number of stock prices in "dataS1" at least equal to S.

\[
\text{dataS3} = \text{Table}[\{\text{dataS2}[[i]], \text{Length}[\text{Select}[\text{dataS1}, \# \geq (\text{dataS2}[[i]]) \&]]\}, \{i, 1, \text{Length}[\text{dataS2}]\}];
\]

The following set, "dataS4", is the logarithmic transformation of "dataS3", and it is graphically depicted as the thick solid curve in what follows: Figure Stock I.

\[
\text{dataS4} = \text{Log}[\text{dataS3}]; \text{ListPlot}[\text{dataS4}, \text{PlotJoined} \to \text{True}, \text{AspectRatio} \to \text{Automatic},
\]

\[
\text{PlotRange} \to \{(0, 11), (0, 8)\}, \text{AxesLabel} \to \{"\log S", "\log N"\},
\]

\[
\text{PlotStyle} \to \text{Thickness}[0.01], \text{PlotLabel} \to "\text{Figure Stock I}"
\]

Utilizing the regression analysis through the least square method, we have the following functional relation:

\[
\log N = 15.2277 - 1.3487 \log S \quad (7)
\]

Note the \(R^2\) value in this regression is relatively high: \(R^2 = 0.936038\).

\text{Regress}[\text{dataS4}, \{1, x\}, x]

\[
\{\text{ParameterTable} \to
\]

\[
\begin{array}{cccc}
\text{Estimate} & \text{SE} & \text{TStat} & \text{PValue} \\
1 & 15.2277 & 0.104317 & 145.976 & 3.648831563119 \times 10^{-443} \\
x & -1.3487 & 0.0150058 & -89.8786 & 9.528049490255 \times 10^{-332}
\end{array}
\]

\text{RSquared} \to 0.936038, \text{AdjustedRSquared} \to 0.935922,

\text{EstimatedVariance} \to 0.0967444, \text{ANOVATable} \to

\[
\begin{array}{ccccc}
\text{DF} & \text{SumOfSq} & \text{MeanSq} & \text{FRatio} & \text{PValue} \\
\text{Model} & 1 & 781.517 & 781.517 & 8078.16 & 9.528049490269 \times 10^{-332} \\
\text{Error} & 552 & 53.4029 & 0.0967444 & \\
\text{Total} & 553 & 834.92
\end{array}
\]

If the stock prices greater than 400 yen but smaller than 8000 are selected, utilizing the regression analysis through the least square method, we have the following functional relation:

\[
\log N = 17.4665 - 1.67308 \log S \quad (8)
\]
Note that the $R^2$ value improves to 0.958507, but the slope deviates further from -1, so that [O1] and [O2] in Section 1 hold in this stock price case.

\[
\text{dataS11} = \text{Select}[\text{dataS1}, 8000 > \# > 400 \&]; \text{dataS21} = \text{Union}[\text{dataS11}];
\text{dataS31} = \text{Table}[\{\text{dataS21}[[i]], \text{Length}[\text{Select}[\text{dataS11}, \# \geq (\text{dataS21}[[i]]) \&]]\}, \{i, 1, \text{Length}[\text{dataS21}]\}]; \text{dataS41} = \text{Log}[\text{dataS31}]; \text{Regress}[\text{dataS41}, \{1, x\}, x]
\]

\[
\begin{array}{cccc}
\text{Estimate} & \text{SE} & \text{TStat} & \text{PValue} \\
1 & 17.4665 & 0.109596 & 159.372 \times 10^{-427} \\
x & -1.67308 & 0.0156619 & -106.825 \times 10^{-343}
\end{array}
\]

\[
\text{RSquared} \rightarrow 0.958507, \text{AdjustedRSquared} \rightarrow 0.958423,
\text{EstimatedVariance} \rightarrow 0.0623308, \text{ANOVATable} \rightarrow
\]

\[
\begin{array}{cccccc}
\text{DF} & \text{SumOfSq} & \text{MeanSq} & \text{FRatio} & \text{PValue} \\
\text{Model} & 1 & 711.29 & 711.29 & 11411.5 & 1.601789283842 \times 10^{-343}
\end{array}
\]

Take an other stock price distribution. On December 30, 1998, there were 1597 stock prices in Tokyo Stock Exchange. The set of all the stock prices is denoted as "dataSA2". The lowest stock price is 34 yen, the highest one is that of "Bank of Japan" stock: 4,650,000 yen. Deleting it from "dataSA2", we obtain "dataSA22". In the same way as above, the logarithmic transformation of "dataSA42" is graphically depicted as the thick solid curve in what follows: Figure Stock II.

\[
\text{dataSA22} = \text{Drop}[\text{dataSA2}, \{1487\}]; \text{dataSA32} = \text{Union}[\text{dataSA22}];
\text{dataSA42} = \text{Table}[\{\text{dataSA32}[[i]], \text{Length}[\text{Select}[\text{dataSA22}, \# \geq (\text{dataSA32}[[i]]) \&]]\}, \{i, 1, \text{Length}[\text{dataSA32}]\}]; \text{dataSA52} = \text{Log}[\text{dataSA42}];
\text{ListPlot}[\text{dataSA52}, \text{PlotJoined} \rightarrow \text{True}, \text{AspectRatio} \rightarrow \text{Automatic},
\text{PlotRange} \rightarrow \{(0, 17), (0, 8)\}, \text{AxesLabel} \rightarrow \{"\text{log } S", \"\text{log } N"\},
\text{PlotLabel} \rightarrow \"\text{Figure Stock II}\" , \text{PlotStyle} \rightarrow \text{Thickness}[0.01]];
\]

Utilizing the regression analysis through the least square method, we have the following functional relation:

\[
\log N = 11.2091 - 0.797215 \log S \quad (9)
\]

Note the high $R^2$ value in this regression: $R^2 = 0.954006$. 

If the stock prices greater than 300 yen, but smaller than 10000 yen, are selected, utilizing the regression analysis through the least square method, we have the following relation:

$$\log N = 12.7308 - 1.01585 \log S$$ \quad (10)

Note that the $R^2$ value improves to 0.993235, and the slope increases, so that [O1] and [O2] in Section 1 hold in this stock price case.

```math
\text{dataSA32A} = \text{Select}[\text{dataSA22}, 10000 > \# > 300 \&]; \text{dataSA32} = \text{Union}[\text{dataSA32A}];
\text{dataSA42} = \text{Table}[\{\text{dataSA32}[\#], \text{Length}[\text{Select}[\text{dataSA22}, \# \in (\text{dataSA32}[\#]) \&]]\},
\{\# , 1, \text{Length}[\text{dataSA32}]\}]; \text{dataSA52} = \text{Log}[\text{dataSA42}]; \text{Regress}[\text{dataSA52}, \{1, x\}, x]
```

<table>
<thead>
<tr>
<th>ParameterTable →</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>x</td>
</tr>
<tr>
<td>RSquared → 0.954006, AdjustedRSquared → 0.953951, EstimatedVariance → 0.060798, ANOVATable →</td>
</tr>
<tr>
<td>DF</td>
</tr>
<tr>
<td>Model</td>
</tr>
<tr>
<td>Error</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

```math
\text{dataSA32} = \text{Union}[\text{dataSA32A}];
\text{dataSA42} = \text{Table}[\{\text{dataSA32}[\#], \text{Length}[\text{Select}[\text{dataSA22}, \# \in (\text{dataSA32}[\#]) \&]]\},
\{\# , 1, \text{Length}[\text{dataSA32}]\}]; \text{dataSA52} = \text{Log}[\text{dataSA42}]; \text{Regress}[\text{dataSA52}, \{1, x\}, x]
```

<table>
<thead>
<tr>
<th>ParameterTable →</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>x</td>
</tr>
<tr>
<td>RSquared → 0.993235, AdjustedRSquared → 0.993223, EstimatedVariance → 0.00495581, ANOVATable →</td>
</tr>
<tr>
<td>DF</td>
</tr>
<tr>
<td>Model</td>
</tr>
<tr>
<td>Error</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
2.3 Reasons for the Non-Linearity

As shown above, in general, the non-linearity holds over the whole region, and the linearity holds on a subset of the whole region, for example, when we include only the larger-sized cities in the data. In this subsection, the reasons for the non-linearity are examined. In the case of stock prices, as the stock price rises due to the excellent corporate performance, the Japanese firms tends to dampen the stock price hike, for example, by distributing free stocks for the firms' stock holders. This tendency may put upper limit for the stock prices, creating the bulge on the diagram. In what follows, the non-linearity in the Japanese city size distribution diagram is examined.

As the population of a city expands, the city area also expands to the suburb area due to the shortage of land or environmental reasons, and the suburban areas may become independent cities. This tendency, however, is common to both the US and Japan, and this cannot explain [O3] in Section 1. As one of the factors, which explain [O3], we may refer to the transfers of the capital in the history of Japan.

Japanese people tends to gather around the capital and/or the cities in which the governments were formed. In the history of Japan, the capital of Japan and/or the cities with governments have been transfered so often, that the previous capital dwindles while the new capital is created. The reasons for the transfer varies from political ones to environmental ones. As an example of the latter case, the transfer of capital from Nara to Kyoto may be mentioned. In the 740s, the Japanese population is estimated as 5.6 million (Kito [1983]). In 740, the capital was transferred from Nara to Kuni, from Kuni to Osaka in 744, from Osaka to Shigaraki in 744, from Shigaraki to Nara in 745. Recently, researchers pointed out that the reason for these transfers might well be the air pollution stemming from the construction of huge statue of Buddha. In this decate, the deeply religious Emperor ordered the construction of a great (gilded copper) statue of Buddha at a temple in Nara. According to the researchers the statute was so large that in plating it with gold by the traditional amalgam method a huge amount of mercury was used (50 tons), and the vaporated mercury must have caused severe air pollution. In those days, the deeply religious Emperor might have regarded the pneumonia and uremia caused by the pollution as a religious curse, and transferred the capital (Miwa [1979], Shirasuga [2004]). In 784, the capital was transferred from Nara to Nagaoka, near Kyoto, then, in 794 from Nagaoka to Kyoto.

As examples of political reasons for the transfer of capital, we may mention the rise of *Samurais*, or warriors. From 794 to 1968, the capital was set at Kyoto by the successive Emperors. In between, however, bestowed the authority from the Emperors, dominant *samurai* clans formed their own governments in different cities. For example, in 1192, with the estimated Japanese population of 6.9 million, a head of feudal lords, *Shogun* Yoritomo formed his government in Kamakura, near Tokyo. Then, in 1590, with the estimated Japanese population of 12.3 million, an other *Samurai*, the chief advisor to the Emperor, Hideyoshi formed his government, in Osaka, and finally in 1603, a *Shogun* Ieyasu formed his government in Edo (Tokyo). The Japanese population increased to 30 million during the 1700s (Kito [1983]). In the Edo Period (1603-1867) 15 *Shoguns* governed Japan with the policy of national isolation. In 1721, the 7th *Shogun* Yoshimune took the national census for the first time. In those days, Edo (Tokyo), with the population of 500,000, was a crowded city, whose population density was much greater than the one of the present-day Tokyo, leading to the high mortality in Edo, which was susceptible to infectious diseases. It is reported that the city-size expansion of Edo was mild (Kito [1983]).

With the end of the political dominance by *Samurais* in 1868, the Emperor transferred the capital from Kyoto to Edo, changing its name as Tokyo. The Japanese population further increased from 43 million in 1900 to 55.9 million in 1920. In 2004, it was 127.68 million. Severe air pollution by SPM prompted the Governer of Tokyo in 2003 to impose stricter environmental legislation than the national one. The Metropolis of Tokyo plans to introduce the *Road Pricing*, following Singapore and London. As mentioned in Introduction, the argument of transferring Japanese capital is also under way for the purpose of the alleviation of environmental problems and other reasons.
3. Two-Stage Simon Process

In this section, following the previous remarks, the *Mathematica* command function of the two-stage Simon Process is constructed, and utilizing it simulations are conducted.

### 3.1 Programming of the Two-Stage Simon Process

Suppose that independently of the growth process of the existing cities with the probability \((1-\pi_1)\) there exists the possibility of the partition of the largest-sized city into small two cities, \(\pi_2\), stemming from the reasons referred to in the previous section. Thus, in this section, the two-stage Simon Process is constructed. In subsection 2.1, it was assumed as a specified case that the base-unit city is created with the probability \(\pi_1=0.1\), while one of the existing cities, say, the one with city size \(i\), expands to the city with city size \(i+1\), with the probability \(1-\pi_1=0.9\). After this stage, suppose that the largest city is partitioned into two smaller-sized cities with the probability \(\pi_2=0.2\). In subsection 2.1, "data10" was the final result. In this section we start from "data9".

```
data9 = Union[{data1[[data7[[1]]]] - {0, 1}, data1[[data7[[2]]]] + {0, 1}}, data8]
{{0, 0}, {1, 7}, {2, 3}, {3, 0}, {4, 1}, {5, 0}}
```

If Random[BernoulliDistribution[0.2]]=0,

```
Random[BernoulliDistribution[0.2]]
0
```

the selection of the elements in "data9", whose number of cities is positive, "data10", say, \({\{1,7\},\{2,3\},\{4,1\}}\), is the city distribution in the next period as in the previous subsection.

```
Select[data9, #[[2]] > 0 &]
{{1, 7}, {2, 3}, {4, 1}}
```

However, if Random[BernoulliDistribution[0.2]]=1, the largest-sized city is partitioned into the two cities with sizes determined at random. Thus, when

```
Random[BernoulliDistribution[0.2]]
1
```
takes place, a different procedure follows as in what follows. First, we define the largest city size as "data11".

```
data10 = Select[data9, #[[2]] > 0 &]; data11 = First[data10][[1]]
4
```

When this city size is partitioned into 2 parts: "data11"=n1+n2, n1 is selected at random by utilizing a *Mathematica* function, "Random", and it is defined as "data12".

```
data12 = Random[Integer, {1, data11 - 1}]
1
```
We add 1 to the number of cities with this city size, and define the resulting set as "data13".

\[
data13 = \text{data9}[\text{data12 + 1, 1}], \text{data9}[\text{data12 + 1, 2}] + 1
\]

\{1, 8\}

The corresponding element in "data9" is replaced by "data13". The resulting set is defined as "data14".

\[
data14 = \text{data9} /. \text{data9}[\text{data12 + 1}] \rightarrow \text{data13}
\]

\{0, 0\}, \{1, 8\}, \{2, 3\}, \{3, 0\}, \{4, 1\}, \{5, 0\}

By assumption, we have n2="data11"-n1, so that we add 1 to the number of the cities with this city size in "data14". The resulting element is defined as "data15".

\[
data15 = \text{data14}[\text{data11 - data12 + 1, 1}], \text{data14}[\text{data11 - data12 + 1, 2}] + 1
\]

\{3, 1\}

The corresponding element in "data14" is replaced by "data15". The resulting set is defined as "data16".

\[
data16 = \text{data14} /. \text{data14}[\text{data11 - data12 + 1}] \rightarrow \text{data15}
\]

\{0, 0\}, \{1, 8\}, \{2, 3\}, \{3, 1\}, \{4, 1\}, \{5, 0\}

In "data16", we must subtract 1 from the number of the largest city, as in what follows. The resulting set is defined as "data17".

\[
data17 = \text{data16} /\ . \ (\text{data16}[\text{data11 + 1}] \rightarrow \text{data16}[\text{data11 + 1}], \text{data16}[\text{data11 + 2}] - 1)
\]

\{0, 0\}, \{1, 8\}, \{2, 3\}, \{3, 1\}, \{4, 0\}, \{5, 0\}

The selection of the elements, in "data17", whose number of cities is positive, say, \{1,8\},{2,3},{3,1}\}, is the city distribution in the next period for the two-stage Simon Process.

\[
Select[\text{data17}, #[[2]] > 0 \&]
\]

\{1, 8\}, \{2, 3\}, \{3, 1\}

The above process is summerized as in what follows.

\[
\text{If[Random[BernoulliDistribution[0.2]] == 0,}
\]

\[
Select[\text{data9}, #[[2]] > 0 \&], \text{data10} = \text{Select[data9, #[[2]] > 0 \&];}
\]

\[
\text{data11 = Last[data10][[1]]}; \text{data12 = Random[Integer, \{1, data11 - 1\}];}
\]

\[
\text{data13 = data9[[data12 + 1, 1]], data9[[data12 + 1, 2]] + 1};
\]

\[
\text{data14 = data9 /. \ (data9[[data12 + 1]]} \rightarrow \text{data13};
\]

\[
\text{data15 = data14[[data11 - data12 + 1, 1]], data14[[data11 - data12 + 1, 2]] + 1};
\]

\[
\text{data16 = data14} /\ . \ (\text{data14[[data11 - data12 + 1]]} \rightarrow \text{data15});
\]

\[
\text{data17 = data16} /\ .
\]

\[
(\text{data16[[data11 + 1]]} \rightarrow \text{data16[[data11 + 1, 1]], data16[[data11 + 1, 2]] - 1});
\]

\[
Select[\text{data17}, #[[2]] > 0 \&]
\]

\{1, 8\}, \{2, 3\}, \{3, 1\}
The collection of all the programs constructed so far leads to the following Mathematica command function, simon3[p1,p2,list], which creates the next period's city size distribution, when the current period's city size distribution, "list", and the probabilities: \( \pi_1 \), and \( \pi_2 \), are given arbitrarily.

```mathematica
simon3[p1_, p2_, list_] :=
Module[{list1, list2, list3, data1, data2, data3, data4, population, data5, data6, data7, data8, data9, data10, data11, data12, data13, data14, data15, data16, data17], list1 = Table[list[[i, 1]], {i, 1, Length[list]}];
list2 = Table[i, {i, 0, Max[list1] + 1}]; list3 = Complement[list2, list1];
data1 = Union[Table[interest[list3[[i]], 0], {i, 1, Length[list3]}], list];
data1a = Select[data1, #[[2]] > 0 &];
f[i_] := Position[data1, Intersection[{data1a[[i]]}, data1][[1]][[1]][[1]]];
scale = Table[data1a[[i, 1]], {i, 1, Length[data1a]}];
density = Table[data1a[[i, 2]], {i, 1, Length[data1a]}];
data2 = Table[data1a[[i, 1]] + data1a[[i, 2]], {i, 1, Length[data1a]}];
data3 = Table[Sum[data2[[i]], {i, 1, k}], {k, 1, Length[data2]}];
data4 = Drop[Prepend[data3, 0] + 1, {Length[data3] + 1}];
population = Last[data3];
data5 =
If[Random[BernoulliDistribution[1 - p1]] > 0, Random[Integer, {1, population}], 0];
data6 = Length[Union[Select[data5 - data4, # > 0 &]]];
data7 = If[data6 > 0, {1, 2}, {f[data6], f[data6 + 1]}];
data8 = Drop[data1, data7];
data9 = Union[{data1[[data7[[1]]]], 0, 1}, data1[[data7[[2]]]] + {0, 1}], data8];
If[Random[BernoulliDistribution[p2]] == 0,
Select[data9, #[[2]] > 0 &], data10 = Select[data9, #[[2]] > 0 &];
data11 = Last[data10][[1]]; data12 = Random[Integer, {1, data11 - 1}];
data13 = {data9[[data12 + 1, 1]], data9[[data12 + 2, 1]] + 1};
data14 = data9/. (data9[data12 + 1] -> data13);
data15 = {data14[[data11 - data12 + 1, 1]], data14[[data11 - data12 + 1, 2]] + 1};
data16 = data14/. (data14[data11 - data12 + 1] -> data15);
data17 = data16/
(data16[[data11 + 1]] -> data16[[data11 + 1, 1]], data16[[data11 + 2, 2]] - 1));
Select[data17, #[[2]] > 0 &]]
```
3.2 Simulations of the Two Stage Simon Process

In this subsection, utilizing the command function constructed in the previous subsection, simulations are attempted. First of all, the exercise mentioned in Fujita, Krugman and Venables [1999, Chapter 12] is reproduced. This is the case, in which \( p_1 = 0.2 \) and \( p_2 = 0 \) for our Mathematica command function, simon3. Setting the initial city size distribution \( C[1] \) as \( \{1, 10\} \), by the Simon Process: \( C[t+1] = \text{simon3}[0.2, 0, C[t]] \), \( t=1, \ldots, T=100,000 \), we have the final city size distribution, \( C[T] = "fdata220" \), and its logarithmic transformation: "fdata220A", is depicted as the thick solid curve in what follows.

Utilizing the regression analysis through the least square method, we have the following functional relation:

\[
\log N = 9.49674 - 1.11646 \log S
\]  
(11)

Note the \( R^2 \) value in this regression is high: \( R^2 = 0.978927 \).

Other simulations, which were not conducted in Fujita, Krugman and Venables [1999, Chapter 12], are attempted in this paper. If \( p_1 = 0.01 \) and \( p_2 = 0 \) are assumed for our command function, simon3, with initial city size distribution, \( C[1] \), given as \( \{1, 10\} \), in the Simon Process: \( C[t+1] = \text{simon3}[0.2, 0, C[t]] \), \( t=1, \ldots, T=100,000 \), we have the final city size distribution, \( C[T] = "fdata001" \), and its logarithmic transformation: "fdata001A", is depicted as the thin solid curve in the above diagram.
Utilizing the regression analysis through the least square method, we have the following functional relation:

\[
\log N = 5.98919 - 0.638409 \log S \quad (12)
\]

Note the relatively high \( R^2 \) value in this regression: \( R^2 = 0.934146 \), while the parameter of Pareto distribution is far from -1.

We examine if it is possible to have the parameter of Pareto distribution close to -1, by suitably selecting a subset of the simulation result. Here, the final city size distribution, "fdata001", with 65 elements, is the one in what follows.

\[
\begin{align*}
\{1,528\}, \{2,168\}, \{3,91\}, \{4,54\}, \{5,24\}, \{6,22\}, \{7,27\}, \{8,15\}, \{9,7\}, \{10,11\}, \{11,12\}, \\
\{12,12\}, \{13,4\}, \{14,7\}, \{15,4\}, \{16,1\}, \{17,1\}, \{18,3\}, \{19,3\}, \{20,1\}, \{21,2\}, \{22,3\}, \\
\{23,1\}, \{24,1\}, \{25,1\}, \{26,1\}, \{27,2\}, \{28,1\}, \{29,3\}, \{30,3\}, \{31,2\}, \{32,1\}, \{35,1\}, \{36,1\}, \\
\{37,1\}, \{38,1\}, \{39,1\}, \{41,1\}, \{43,1\}, \{48,2\}, \{51,1\}, \{52,1\}, \{54,1\}, \{60,1\}, \{61,1\}, \\
\{64,1\}, \{65,1\}, \{67,2\}, \{69,1\}, \{73,1\}, \{75,1\}, \{138,1\}, \{153,1\}, \{227,1\}, \{253,1\}, \{276,1\}, \\
\{289,1\}, \{445,1\}, \{480,1\}, \{836,1\}, \{851,1\}, \{962,1\}, \{4360,1\}, \{13812,1\}, \{72500,1\}
\end{align*}
\]

With the high probability of expansion in the existing cities: \( 1 - \pi_1 = 0.99 \), the largest city size is 72500 in the simulation, which is extremely large. Thus, if the smallest 60 city sizes are selected with extremely large cities deleted, then, we have the following regression result:

\[
\log N = 6.90396 - 0.982983 \log S \quad (13)
\]

Note the \( R^2 \) value in this regression is high: \( R^2 = 0.991358 \).

\begin{verbatim}
Regress[fdata001AA, {1, x}, x]

<table>
<thead>
<tr>
<th>Estimate</th>
<th>SE</th>
<th>TStat</th>
<th>PValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>ParameterTable 1</td>
<td>6.90396</td>
<td>0.044566</td>
<td>154.915</td>
</tr>
<tr>
<td>x</td>
<td>-0.982983</td>
<td>0.0120507</td>
<td>-81.5703</td>
</tr>
</tbody>
</table>

RSquared → 0.991358, AdjustedRSquared → 0.991209, EstimatedVariance → 0.0148411,

<table>
<thead>
<tr>
<th>DF</th>
<th>SumOfSq</th>
<th>MeanSq</th>
<th>FRatio</th>
<th>PValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1</td>
<td>98.7484</td>
<td>98.7484</td>
<td>6653.71</td>
</tr>
<tr>
<td>Error</td>
<td>58</td>
<td>0.860784</td>
<td>0.0148411</td>
<td>0.</td>
</tr>
<tr>
<td>Total</td>
<td>59</td>
<td>99.6092</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\end{verbatim}

If \( \pi_1 = 0.70 \) and \( \pi_2 = 0 \) are assumed for our command function, simon3, with initial city size distribution \( C[1] \) as \( \{1,10\} \), in the Simon Process: \( C[t+1]=\text{simon3}[0.2, 0, C[t]] \), \( t=1, ..., T=100,000 \), we have the final city size distribution, \( C[T]="fdata070" \), with 34 elements, and its logarithmic transformation: "fdata070AA", is depicted as the thin solid curve in the above diagram.

Utilizing the regression analysis through the least square method, we have the following functional relation:

\[
\log N = 10.8689 - 2.42998 \log S \quad (14)
\]

Note the relatively high \( R^2 \) value in this regression: \( R^2 = 0.966466 \), while the parameter of Pareto distribution is far from -1. We examine if it is possible to have the parameter of Pareto distribution close to -1, by suitably selecting a subset of the simulation result. Here, the final city size distribution, "fdata070", is given by the following.

\[
\begin{align*}
\{1,53910\}, \{2,10124\}, \{3,3201\}, \{4,1293\}, \{5,604\}, \{6,316\}, \{7,205\}, \{8,129\}, \{9,81\}, \\
\{10,53\}, \{11,41\}, \{12,27\}, \{13,18\}, \{14,7\}, \{15,12\}, \{16,10\}, \{17,7\}, \{18,3\}, \{19,4\},
\end{align*}
\]
With the low probability of expansion in the existing cities: $1 - \pi_1 = 0.30$, the number of the base-unit city is quite large in the simulation result. Thus, if the largest 20 city sizes are selected with small cities deleted, then, we have the following regression result:

$$\log N = 7.50416 - 1.50796 \log S$$ (15)

Note the relatively high $R^2$ value in this regression: $R^2 = 0.940985$.

The conclusion in the simulations so far are as follows.

[S1] As $\pi_1$ increases with $\pi_2$ set at 0, the overall slope increases.

[S2] With the deletion of the extreme cities, however, the slope is close to 1 in all cases.

Next, consider the cases with positive $\pi_2$. First of all, it is examined if the upper bound on the city size exists when the positiveness of $\pi_2$ is assumed. If $\pi_2$ is small, then it is expected that the upper bound on the city size might not exist. Thus, suppose that $\pi_2 = 0.01$, to check the existence of the upper bound. Furthermore, suppose that $\pi_1 = 0.01$: high possibility of expansion in the existing cities.

i. With the initial city size distribution, $C[1]$ as $\{1,1\}$, in the Simon Process: $C[t+1]=\text{simon3}[0.01, 0.01, C[t]]$, $t=1, \ldots, T=1,000,000$, we have the final city size distribution, $C[T]$, with 194 the largest city size, as in what follows:

$${\{1,5014\}, \{2,1733\}, \{3,833\}, \{4,541\}, \{5,393\}, \{6,263\}, \{7,198\}, \{8,202\}, \{9,170\}, \{10,157\}, \{11,137\}, \{12,113\}, \{13,99\}, \{14,102\}, \{15,105\}, \ldots, \{170,55\}, \{171,58\}, \{172,46\}, \{173,52\}, \{174,41\}, \{175,36\}, \{176,50\}, \{177,54\}, \{178,54\}, \{179,51\}, \{180,60\}, \{181,55\}, \{182,48\}, \{183,45\}, \{184,63\}, \{185,56\}, \{186,53\}, \{187,50\}, \{188,57\}, \{189,51\}, \{190,43\}, \{191,42\}, \{192,52\}, \{193,17\}, \{194,3\}}.$$

ii. With the initial city size distribution, $C[1]$ as $\{1,1\}, \{2,1\}, \ldots, \{200,1\}$, in the Simon Process: $C[t+1]=\text{simon3}[0.01, 0.01, C[t]]$, $t=1, \ldots, T=1,000,000$, we have the final city size distribution, $C[T]$, with 194 the largest city size, as in what follows:

$${\{1,5149\}, \{2,1708\}, \{3,933\}, \{4,516\}, \{5,344\}, \{6,287\}, \{7,218\}, \{8,191\}, \{9,183\}, \{10,130\}, \{11,133\}, \{12,111\}, \{13,114\}, \{14,79\}, \{15,103\}, \ldots, \{170,47\}, \{171,60\}, \{172,48\}, \{173,47\}, \{174,60\}, \{175,44\}, \{176,60\}, \{177,54\}, \{178,53\}, \{179,48\}, \{180,42\}, \{181,54\}, \{182,56\}, \{183,51\}, \{184,60\}, \{185,49\}, \{186,61\}, \{187,37\}, \{188,57\}, \{189,45\}, \{190,54\}, \{191,42\}, \{192,52\}, \{193,17\}, \{194,3\}}.$$

With the initial city size distribution, $C[1]$ as $\{1, 1\}, \{2, 1\}, ..., \{500, 1\}$, in the Simon Process: $C[t+1]=\text{simon3}[0.01, 0.01, C[t]]$, $t=1, ..., T=1,000,000$, we have the final city size distribution, $C[T]$, with 211 the largest city size, as in what follows:

$\{1, 15732\}, \{2, 1863\}, \{3, 914\}, \{4, 581\}, \{5, 374\}, \{6, 291\}, \{7, 232\}, \{8, 157\}, \{9, 150\}, \{10, 109\}, \{11, 106\}, \{12, 91\}, \{13, 83\}, \{14, 76\}, \{15, 80\}, ..., \{170, 60\}, \{171, 38\}, \{172, 33\}, \{173, 50\}, \{174, 53\}, \{175, 50\}, \{176, 43\}, \{177, 50\}, \{178, 44\}, \{179, 48\}, \{180, 53\}, \{181, 55\}, \{182, 53\}, \{183, 55\}, \{184, 39\}, \{185, 57\}, \{186, 49\}, \{187, 42\}, \{188, 41\}, \{189, 47\}, \{190, 58\}, \{191, 46\}, \{192, 54\}, \{193, 42\}, \{194, 49\}, \{195, 45\}, \{196, 59\}, \{197, 49\}, \{198, 42\}, \{199, 55\}, \{200, 61\}, \{201, 51\}, \{202, 60\}, \{203, 34\}, \{204, 44\}, \{205, 53\}, \{206, 61\}, \{207, 51\}, \{208, 60\}, \{209, 53\}, \{210, 13\}, \{211, 1\}$.

iv. With the initial city size distribution, $C[1]$ as $\{1, 1\}, \{2, 1\}, ..., \{500, 1\}$, in the Simon Process: $C[t+1]=\text{simon3}[0.01, 0.01, C[t]]$, $t=1, ..., T=200,000$, the last element of the final city size distribution was $\{256, 10\}$. Modifying $T$, from 200,000 to 400,000, 600,000, 800,000, 1,000,000, and finally to 1,200,000, the corresponding last elements of the final city size distributions, constructed in this way, were $\{256, 10\}, \{229, 1\}, \{217, 18\}, \{214, 13\}, \{211, 1\}$, and $\{209, 23\}$, respectively. So far, there is a tendency to have the upper bound on the largest city-size, when $p_1=p_2=0.01$ is assumed.

v. What would happen when $p_1$ is large and $p_2=0.01$. With the initial city size distribution, $C[1]$ as $\{1, 1\}$, in the Simon Process: $C[t+1]=\text{simon3}[0.7, 0.01, C[t]]$, $t=1, ..., T=200,000$, we have the final city size distribution, $C[T]$, with 7 the largest city size, as in what follows:

$\{1, 107879\}, \{2, 20669\}, \{3, 6970\}, \{4, 3173\}, \{5, 1716\}, \{6, 1142\}, \{7, 7250\}$.

Meanwhile, with the initial city size distribution, $C[1]$ as $\{1, 1\}, \{2, 1\}, ..., \{20, 1\}$, in the Simon Process: $C[t+1]=\text{simon3}[0.7, 0.01, C[t]]$, $t=1, ..., T=200,000$, we have the final city size distribution, $C[T]$, with 7 the largest city size, as in what follows:

$\{1, 108434\}, \{2, 20852\}, \{3, 6684\}, \{4, 3092\}, \{5, 1647\}, \{6, 1139\}, \{7, 369\}$.

There is a tendency to have the upper bound on the largest city-size, when $p_1=0.7$ and $p_2=0.01$ are assumed. From these simulations, we may conclude the following.

[S3] When $p_2>0$, there exists the upper bound on the city size at final city size distribution, $C[T]$.

We might even conclude that the final city size distribution is convergent when $p_2>0$. Gabaix [1997] examined the convergence of the Simon process when the largest city size is fixed at 64.
Next, let us proceed to the concavity. Suppose that \( \pi_1=0.2 \) and \( \pi_2=0.01 \). Setting the initial city size distribution, \( C[1] \) as \( \{(1,10)\} \), by the Simon Process: \( C[t+1]=\text{simon3}[0.2, 0.01, C[t]], t=1, \ldots, T=100,000 \), we have the final city size distribution, \( C[T]=\text{"fdata20"} \), and its logarithmic transformation: \( \text{"fdata20AA"} \), is depicted as the thick dashed curve in what follows, while the thick solid curve is the result with \( \pi_1=0.2 \) and \( \pi_2=0 \). When \( \pi_1=0.2 \) and \( \pi_2=0.2 \), the final city size distribution is depicted as the dashed curve with a broad interval. In both cases, the concavity appears when \( \pi_2 > 0 \).

Utilizing the regression analysis on \( \text{"fdata20AA"} \), we have the following functional relation:

\[
\log N = 10.7013 - 1.29716 \log S \quad (16)
\]

Note, however, the low \( R^2 \) value in this regression: \( R^2 = 0.71864 \).

Other exercises are conducted with different parameters. Suppose that \( \pi_1=0.7 \) and \( \pi_2=0.01 \). Setting the initial city size distribution, \( C[1] \) as \( \{(1,10)\} \), by the Simon Process: \( C[t+1]=\text{simon3}[0.7, 0.01, C[t]], t=1, \ldots, T=100,000 \), we have the final city size distribution, \( C[T]=\text{"fdataA001"} \), and its logarithmic transformation is depicted as the dashed curve in what follows, while the thin solid curve is the result with \( \pi_1=0.7 \) and \( \pi_2=0 \). When \( \pi_1=0.7 \) and \( \pi_2=0.1 \), the final city size distribution is depicted as the dashed curve with a broad interval. In both cases, the concavity appears when \( \pi_2 > 0 \).
Thus, we may conclude the following.

[S4] Compared with when \( p_2 = 0 \), the stronger concavity of the functional relation between \( \log S \) and \( \log N \) emerges when \( p_2 > 0 \), while \( R^2 \) is lower.

Conclusions

Zipf [1949] noticed that the US city size distribution exhibits the Pareto distribution, in which \( \log N \) is a linear function of \( \log S \) with slope 1, where \( S \) is the city size: the city's population, and \( N \) is the number of cities whose population is at least equal to \( S \). Fujita, Krugman and Venables [1999] pointed out that the Simon process can create the US city size distribution when the probability of the emergence of the base-unit city is 0.2. The Simon process is a non-economic growth model of "lumps" with the probability \( p_1 \) of the emergence of base-unit lump, while \( 1 - p_1 \) is the probability of the expansion of the existing lumps. To be precise, supposing that "lump" is a city, they pointed out that when \( p_1 = 0.2 \) the Simon process can create the US city size distribution. In this paper, we started with pointing out the fact that the Japanese city size distribution exhibits more concavity in the functional relation between \( \log S \) and \( \log N \), than the one in the US, when not only the larger-sized cities but also the smaller-sized cities are included in the data set. The main aim of this paper is to examine if the Japanese city size distribution can be created by modifying Simon process. In order to do this, somewhat more general Simon process was proposed. Along with the probability of \( p_1 \) we propose another probability \( p_2 \) which is the probability of the break-up of the largest city into two smaller cities. Thus, the two-stage Simon process was proposed and examined. The reason of this extension of Simon process consists in the Japanese history. By political and environmental reasons Japanese capitals and/or the cities in which the government was established were transferred: the break-up of the largest city.

In Section 1 of this paper, the difference between the US city size distribution and the Japanese one was examined and the above observation was pointed out. In Section 2, after the Mathematica programming of the traditional Simon was explained, the Japanese history was discussed, as well as other examples of Pareto distribution in terms of the Japanese stock price distribution. In Section 3, Mathematica programming of the command function for the two-stage Simon process was attempted for the simulations in terms of this command function. First of all, assuming \( p_2 = 0 \): the traditional Simon process, simulations of other cases than the one mentioned in Fujita, Krugman and Venables [1999] were conducted, with the conclusion that as \( p_1 \) increases the slope of the Pareto distribution increases. Also, it was found that when \( p_1 \) is small, the slope is close to 1 when only the smaller-sized cities are selected for the data set in the regression, while when \( p_1 \) is large, the slope is close to 1 when only the larger-sized cities are selected for the data set in the regression. When \( p_2 > 0 \), it was found that the there exists a tendency of the existence of upper-bound on the city size. As expected, the functional relation between \( \log S \) and \( \log N \) is concave when \( p_2 > 0 \). However, the linearity of the functional relation between \( \log S \) and \( \log N \) for a subset of the data set was not guaranteed when \( p_2 > 0 \). In this sense, more sophisticated mechanism is required to create the Japanese city size distribution by the Simon process. This task will be pursued in the subsequent papers.
References


Simon, H. [1955], "On a Class of Skew Distribution Functions," *Biometrika* 42.
