On the Gibrat Distribution
Author(s): M. Kalecki
Source: *Econometrica*, Vol. 13, No. 2 (Apr., 1945), pp. 161-170
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/1907013
Accessed: 23/04/2013 14:29

ON THE GIBRAT DISTRIBUTION[1]

By M. KALECKI

THE GIBRAT APPROACH

1. IT WAS a great achievement of Gibrat[2] to show that the distribution of the *logarithms* of some economic variates (for instance, the distribution of factories according to the number of workers) is approximately normal. The explanation of this phenomenon by Gibrat may be presented in a rigorous form as follows:

Let us denote the variate X (for instance the number of workers in a factory) at a certain date by X_0. Let us further assume that subsequently it undergoes a series of random independent *proportionate* changes $m_1, m_2, \cdots, m_n$ (Gibrat's *loi de l'effet proportionnel*).[3] Thus at the end of the period in which these changes have taken place the value of the variate will have become $X_0(1+m_1)(1+m_2) \cdots (1+m_n)$ and its natural logarithm $= \log X_0 + \log(1+m_1) + \log(1+m_2) + \cdots + \log(1+m_n)$. If we denote the deviation from the mean of $\log X_0$ by Y_0 and the deviation from the mean of $\log(1+m_k)$ by y_k, the deviation from the mean of this expression becomes $Y_0+y_1+y_2+\cdots+y_n$. The absolute value of m_k may be assumed small as compared with 1. It follows that the absolute value of $\log(1+m_k)$ and consequently that of y_k is also small as compared with 1. As the second moment of $y_1+y_2+\cdots+y_n$ is equal to the sum of the second moments of $y_1, y_2, \cdots, y_n$, it may be assumed that if n is sufficiently large the standard deviation of $y_1+y_2+\cdots+y_n$ is equal to or greater than 1 (provided the standard deviation of y_n does not fall below a certain level as n increases.) Thus y_k is small as compared with the standard deviation of $y_1+y_2+\cdots+y_n$. With this condition fulfilled the distribution of $y_1+y_2+\cdots+y_n$ is approximately normal (according to the Laplace-Liapounoff theorem[4]).

Further if n is so large that the standard deviation of $y_1+y_2+\cdots+y_n$ is large as compared with the standard deviation of Y_0 also, the distribution of $Y_0+y_1+y_2+\cdots+y_n$ will not differ much from normality. Whatever the distribution of Y at the initial date, with the lapse of time it approaches normality more and more.

2. This argument is formally correct but it may be shown that its

[1] I am much indebted to Mr. D. G. Champernowne for his comments which enabled me to improve upon the first draft of this article.

[2] R. Gibrat, *Les inégalités économiques*, Paris, Sirey, 1931, 296 pp.

[3] Actually the "law of proportionate effect" has been known long before Gibrat.

[4] See, for instance, J. V. Uspensky, *Introduction to Mathematical Probability*, New York, McGraw-Hill, 1937, 411 pp., esp. pp. 291–296.

implications are unrealistic and that consequently there must be something wrong in the underlying assumptions. Indeed, the argument implies that as time goes by the standard deviation of the logarithm of the variate considered increases continuously. In the case of many economic phenomena, however, no tendency for such an increase is apparent (for instance in distribution of incomes). And also for a priori reasons it is clear that changes in the standard deviation of the logarithm of a given variate are to a great extent determined by *economic* forces. It follows therefore that the standard deviation of the variate Y is fully or partly constrained and as a result the random changes are not independent of the values of the variate Y that are subject to them. It may, however, be demonstrated that on certain conditions also in this case the distribution tends towards normality as the time goes by.

EVOLUTION OF FREQUENCY CURVES

3. Let us consider first a case where the second moment M of Y is kept constant throughout time by economic forces. This means that if Y is subject to a change by y,

$$\frac{1}{N}\sum (Y+y)^2 = \frac{1}{N}\sum Y^2 = M. \tag{1}$$

We thus have:

$$\sum Y^2 + 2\sum Yy + \sum y^2 = \sum Y^2$$

or

$$2\sum Yy = -\sum y^2. \tag{2}$$

It follows that there is a negative correlation between Y and y. Let us assume that the correlation is of the type:

$$y = -\alpha Y + z \tag{3}$$

where α is a constant and z is independent of Y. If we substitute this expression for y into the left-hand side of the equation (2) we obtain:

$$-2\alpha\sum Y^2 + 2\sum Yz = -\sum y^2.$$

Since z is independent of Y

$$2\sum Yz = 0$$

and

$$\alpha = \frac{\sum y^2}{2\sum Y^2}. \tag{4}$$

$\sum y^2/N$, which is the second moment of y, may be assumed small as compared with M. Thus α is small as compared with 1. From this and from α's being positive it follows that:

$$0 < 1 - \alpha < 1. \tag{5}$$

We have further:

$$Y + y = (1 - \alpha)Y + z. \tag{6}$$

It follows that:

$$Y_0 + y_1 = Y_0(1 - \alpha_1) + z_1,$$
$$Y_0 + y_1 + y_2 = Y_0(1 - \alpha_1)(1 - \alpha_2) + z_1(1 - \alpha_1) + z_2,$$

and, finally,

$$\begin{aligned} Y_0 + y_1 + y_2 + \cdots + y_n = {} & Y_0(1 - \alpha_1)(1 - \alpha_2) \cdots (1 - \alpha_n) \\ & + z_1(1 - \alpha_2) \cdots (1 - \alpha_n) \cdots \\ & + z_{n-1}(1 - \alpha_n) + z_n. \end{aligned} \tag{7}$$

Since $0<1-\alpha_k<1$ the absolute value of the first component is small as compared with $\sqrt{\sum Y_0{}^2/N}=\sqrt{M}$ if n is sufficiently large, provided α_n does not tend to zero when n tends to ∞. The same is true of other components if the absolute value of z_k is assumed small as compared with M.

As $\sqrt{M}$ is the standard deviation of $Y_0+y_1+y_2+\cdots+y_n$ (M being kept constant throughout) and the z_k are independent of Y_0 and each other, we may conclude that, if n is sufficiently large, the distribution of $Y_0+y_1+y_2+\cdots+y_n$ is approximately normal.[5]

4. We shall now consider the general case where the second moment M of Y changes through time. This may happen in three ways: (a) The second moment increases merely under the influence of random shocks; this case was considered in the section on the "Gibrat Approach." (b) The change in M is fully determined by economic forces. (c) The influence of economic forces upon M is not so rigid as to prevent it fully from being influenced by shocks. We can say that in this case—intermediate between (a) and (b)—M is semi-constrained. However, the difference between (b) and (c) does not affect our subsequent argument. We shall, however, assume throughout it that the change of the second moment of Y, which we denote by ΔM, is of the same order of magnitude as the second moment of y, i.e., $\sum y^2/N$. As

[5] Mr. Champernowne points out that this result is confirmed, as a special case, by the results outlined in his paper read at the meeting of the Econometric Society in Oxford in 1936 and reported in ECONOMETRICA, Vol. 5, October, 1937, pp. 379–381.

$\sum Y^2/N = M$ it follows that $\Delta M/M$ is a small number of the same order as $\alpha = \sum y^2/2\sum Y^2$.

Imagine that the variate Y has been subjected to a change by y and at the same time M has changed by ΔM. We shall have the relation

$$(1') \qquad \frac{1}{N}\sum (Y+y)^2 = M + \Delta M = \frac{1}{N}\sum Y^2 + \Delta M.$$

It follows that:

$$(2') \qquad 2\sum Yy = -\sum Y^2 + N\cdot\Delta M,$$

and, as in the case previously considered, if we assume the correlation to be linear, it follows that

$$(3') \qquad y = -\beta Y + z$$

where z is independent of Y and β is determined by the equation

$$(4') \qquad \beta = \frac{\sum y^2 - N\cdot\Delta M}{2\sum Y^2} = \alpha - \frac{\Delta M}{2M}$$

and

$$1 - \beta = 1 + \frac{\Delta M}{2M} - \alpha.$$

Let us divide both sides of this equation by $\sqrt{1+\Delta M/M}$ and denote

$$\frac{1-\beta}{\sqrt{1+\frac{\Delta M}{M}}}$$

by $1-\gamma$. We have

$$1-\gamma = \frac{1+\frac{\Delta M}{2M}}{\sqrt{1+\frac{\Delta M}{M}}} - \frac{\alpha}{\sqrt{1+\frac{\Delta M}{M}}}.$$

$\Delta M/M$ being a small quantity of the same order as α and α being positive, it follows that γ also is of the same order as α and is positive, for

$$\frac{1+\frac{\Delta M}{2M}}{\sqrt{1+\frac{\Delta M}{M}}}$$

differs from 1 by a small number of the second order and the same is true of the difference between

$$\frac{\alpha}{\sqrt{1+\frac{\Delta M}{M}}}$$

and α. We have thus

$$1-\beta=(1-\gamma)\sqrt{\frac{M+\Delta M}{M}} \tag{8}$$

and

$$0<\gamma<1, \tag{5'}$$

where γ is of the same order as α.

Further:

$$Y+y=(1-\beta)Y+z, \tag{6'}$$

from which follows

$$\begin{aligned} Y_0+y_1+y_2+\cdots+y_n = {} & Y_0(1-\beta_1)(1-\beta_2)\cdots(1-\beta_n) \\ & +z_1(1-\beta_2)\cdots(1-\beta_n)+\cdots \\ & +z_{n-1}(1-\beta_n)+z_n. \end{aligned} \tag{7'}$$

Let us denote the second moment of Y_0 by M_0, that of Y_0+y_1 by M_1, that of $Y_0+y_1+y_2$ by M_2, etc. We then have from the equation (8)

$$1-\beta_k=(1-\gamma_k)\sqrt{\frac{M_k}{M_{k-1}}},$$

where γ_k fulfils the condition

$$0<1-\gamma_k<1$$

and is of the same order as $\alpha_k=\sum y_k^2/N$. It follows that

$$(1-\beta_1)(1-\beta_2)\cdots(1-\beta_n)=(1-\gamma_1)(1-\gamma_2)\cdots(1-\gamma_n)\sqrt{\frac{M_n}{M_0}},$$

$$(1-\beta_2)(1-\beta_3)\cdots(1-\beta_n)=(1-\gamma_2)(1-\gamma_3)\cdots(1-\gamma_n)\sqrt{\frac{M_n}{M_1}},$$

and so on. Thus equation (7′) may now be written

$$\begin{aligned} & Y_0+y_1+y_2+\cdots+y_n \\ & \qquad = \frac{Y_0}{\sqrt{M_0}}(1-\gamma_1)\cdots(1-\gamma_n)\sqrt{M_n} \end{aligned}$$

$$
(7'') \qquad + \frac{z_1}{\sqrt{M_1}} (1 - \gamma_2) \cdots (1 - \gamma_n)\sqrt{M_n} + \cdots
$$

$$
+ \frac{z_{n-1}}{\sqrt{M_{n-1}}} (1 - \gamma_n)\sqrt{M_n} + z_n.
$$

Since $0<1-\gamma_k<1$, the absolute value of the first component of the right-hand side of equation (7″) becomes small as compared with M_n, i.e., with the standard deviation of $Y_0+y_1+y_2+ \cdots +y_n$, if n is sufficiently large provided α_n does not tend to zero when n tends to ∞. The same is true of other components if the absolute values of z_k are assumed to be small as compared with $\sqrt{M_k}$, i.e., with the standard deviation of $Y_0+y_1+y_2+ \cdots +y_k$. Taking into account that the z_k are independent of Y_0 and of each other, we can conclude that if n is sufficiently large the distribution of $Y_0+y_1+y_2+ \cdots +y_n$ is approximately normal.

5. So far we have assumed that the correlation between y and Y is of the form:

$$
(3') \qquad y = - \beta Y + z,
$$

where z is independent of Y. We can make a more general assumption. Let us denote a function of Y by U and the increase in U corresponding to y by u. u is correlated with U because y is correlated with Y. Let us assume that for some function U there is fulfilled the condition:

$$
(3'') \qquad u = - \lambda U + v,
$$

where λ is a constant and v is independent of U.[6] If we further denote the respective deviations from the mean by u', U', and v', we have

$$
(3''') \qquad u' = - \lambda U' + v'.
$$

It follows from the above that, if U is subjected to a long series of

[6] This condition need not necessarily be fulfilled. The relation between y and Y can be put in its general form:

$$
y = F(Y, t),
$$

where t is independent of Y. As t is small this amounts to:

$$
y = \phi(Y) + \Psi(Y)\cdot t.
$$

It follows that:

$$
\frac{u}{\dfrac{dU}{dY}} = \phi(Y) + \Psi(Y)\cdot t \qquad \text{or} \qquad u = \frac{dU}{dY}\phi(Y) + \frac{dU}{dY}\Psi(Y)\cdot t.
$$

To make the second component independent of Y and thus of U, the latter must be chosen so as to render $(dU/dY)\Psi(Y)$ a constant. But this choice of the function U will in general not render $(dU/dY)\phi(Y)$ a linear function of U and therefore in general the condition $u=-\lambda U+v$ is *not* fulfilled.

shocks u, the distribution of U will approach normality provided the change in the second moment of U' is throughout of the same order as $\sum u'^2/N$ (cf. p. 163).

Let us now remember that by Y we denoted the logarithm of the deviation of log X from the mean. Thus if we denote by G the geometrical average of X we have

$$Y = \log \frac{X}{G} \cdot$$

Further, U being a function of Y, we can write

$$U = f\left(\frac{X}{G}\right).$$

Thus if the condition $u = -\lambda U + v$ is fulfilled a series of shocks leads to a normal distribution in terms of $f(X/G)$. If $U = Y$, i.e., if $f(X/G) = \log\,(X/G)$, we have then the case initially considered where the series of shocks leads to a normal distribution of log X.

THE GIBRAT PHENOMENON

6. We shall now illustrate the Gibrat phenomenon on two examples. The first is the distribution of factories according to the number of workers in the U. S. A. This follows the pattern of a normal distribution of the logarithms of the variate which may be called the simple Gibrat phenomenon. The second is the distribution of incomes in the United Kingdom. In the latter example we shall have to do with a modified Gibrat phenomenon for which Gibrat was unable to give a reasonable theoretical interpretation. We shall make an attempt to fill this gap by postulating an appropriate shape of the function $U = f(X/G)$.

In Table 1 is given (a) the actual distribution of factories according to the number of workers in the United States manufacturing industry in 1937; (b) the figures n calculated from the formulae:

$$n = \frac{N}{\sqrt{2\pi}} \int_w^\infty e^{-w^2/2} dw,$$

$$w = 1.20 \log X - 1.091,$$

where N is the total number of factories $= 166.8$ thousand and X the number of workers in factory; throughout the paper below log stands for common or decimal logarithm.

The agreement between the actual and calculated series is fairly good: the distribution of logarithms of the number of workers in a factory is approximately normal. This may be explained by assuming

TABLE 1

DISTRIBUTION OF FACTORIES IN THE U. S. A. MANUFACTURING INDUSTRY IN 1937

Number of workers X	Number of factories n Actual (thousands)	Calculated (thousands)	Percentage error
Above 0	166.8	166.8	0 %
Above 5	97.8	100.1	+2.4
Above 20	51.4	53.2	+3.5
Above 50	28.3	28.6	+1.1
Above 100	16.4	15.9	−3.0
Above 250	6.55	6.17	−5.8
Above 500	2.64	2.65	+0.4
Above 1,000	0.978	1.009	+3.2
Above 2,000	0.241	0.235	−2.5

that the condition $y=-\beta Y+z$, where z is independent of Y, is fulfilled (cf. pp. 163 ff).

7. We shall now consider the distribution of incomes in the U. K. assessed in 1938/39. The actual figures are given in the second column of Table 2.[7] Let us first try to approximate this distribution by a normal distribution of log X as in our first example. The total number of incomes in 1938/39—exclusive of unemployed on the dole and old pensioners, and considering the income of husband and wife jointly as one income to fit the Inland Revenue statistics—may be estimated as being of the order of 20 million. The calculation made for $N=20$ million shows that the actual distribution is not even approximately normal in terms of log X.

But another phenomenon may be observed. If we consider the normal distribution not of log X but of log $(X-A)$ where A is an appropriately chosen income, we obtain a very good approximation to the actual distribution. In Table 2 in addition to the actual distribution above £250 are given figures calculated from the formulae:

$$n = \frac{N}{\sqrt{2\pi}} \int_w^\infty e^{-w^2/2} dw,$$

$$w = 1.10 \log (X - 155) - 1.042,$$

where $N=20$ million.

[7] Source: Statement by Captain Crookshank in the House of Commons July 23, 1942. The statement contains also the number of incomes from £125 to £250. When examined in the light of the data on aggregate personal income this figure appears much too small. This may be explained by the fact that many income earners in this group, who would not in any case pay taxes because of the operation of allowances, have not submitted their income declarations.

This formula has been fitted to the actual n over the range from £250 to £10,000, because the number of incomes above £25,000, £50,000, and £100,000 given in the source referred to are very unprecise on account of rounding off. However, it gives a fairly good approximation also for the number of incomes above £25,000 and £50,000, only for number of incomes above £100,000 the percentage error is rather high.

But although the "modified Gibrat curve" gives a good approximation of the actual distribution of incomes above £250 it is easy to see that such is not the case if we take into consideration *all* incomes. For it follows directly from the above formulae that the number of incomes over £155 is equal to N, i.e., to the total number of incomes; or in other words that all incomes are above £155 p.a., which is, of

TABLE 2

DISTRIBUTION OF PERSONAL INCOMES IN THE U. K. ASSESSED IN 1938/39

Income p.a. X	Number of Incomes n (thousands) Actual	Number of Incomes n (thousands) Calculated	Percentage Error
Above £ 250	2,550	2,568	+ 0.7%
Above 500	800	802	+ 0.2
Above 1,000	300	294	− 2.0
Above 2,000	105	107.4	+ 2.3
Above 3,000	58.5	58.4	− 0.2
Above 5,000	26.5	26.0	− 1.9
Above 10,000	8.0	8.1	+ 1.2
Above 25,000	1.4	1.49	+ 6.5
Above 50,000	0.4	0.37	− 7.5
Above 100,000	0.1	0.083	−17.0

course, fantastic. The "modified Gibrat curve" thus provides a good approximation to the actual distribution only for a limited range of incomes. We shall see that as such it may be easily interpreted in the light of results arrived at in the preceding section.

8. Let us postulate that the function $U=f(X/G)$ fulfilling approximately the condition (cf. p. 167)

$$u = -\lambda U + v \qquad (3'')$$

is of such a shape that for $X/G \geqq b$ it nearly coincides with $\log(X/G-a)$ where $a<b$. It follows from the argument on p. 166 that a long series of shocks leads to a distribution that is normal in terms of $f(X/G)$. We thus have:

$$n = \frac{N}{\sqrt{2\pi}} \int_{w}^{\infty} e^{-w^2/2} dw,$$

$$w = p f\left(\frac{X}{G}\right) + q.$$

But for $X/G < b$ according to our assumption $f(X/G)$ is approximately equal to $\log (X/G - a)$. We thus have approximately for $X/G \geqq b$:

$$w = p \log \left(\frac{X}{G} - a\right) + q.$$

If we denote Gb by B and Ga by A and $q - \log G$ by q', we can say that for $X \geqq B$ we have

$$w = p \log (X - A) + q',$$

where $A < B$. We thus obtain a "modified Gibrat distribution" for incomes higher than (or equal to) B. For large X this distribution approaches normality in terms of $\log X$ because $\log X - \log (X - A)$ tends to zero when X tends to ∞.

It should be stressed that this "explanation" provides us only with a *model* of the process leading to the distribution considered. For: (a) The condition $u = -\lambda U + v$ where v is independent of U need not be fulfilled in actual fact (cf. footnote to p. 166). (b) Even if this condition is fulfilled the function $U = f(X/G)$ satisfying it may change in time, whereas our argument is based on the assumption that its shape is unaltered throughout the process.

Oxford Institute of Statistics