Mathematical Social Sciences 18 (1989) 99-105
North-Holland

NOTE

THE OPTIMUM SIZE DISTRIBUTION OF FIRMS*

Satya R. CHAKRAVARTY
Indian Statistical Institute, Calcutta, India and Institute für Wirtschaftstheorie und Operations Research, University of Karlsruhe, D-7500 Karlsruhe, F.R.G.

Communicated by F.W. Roush
Received 11 July 1988
Revised 22 November 1988

An industry performance evaluation function, which is used for ranking alternative output distributions, is usually expressed as a trading-off of the total output and a 'numbers equivalent', an inverse index of industrial concentration. Using the Shannon entropy numbers equivalent, this paper determines the output distribution that maximises the industry performance evaluation function subject to the constraints that (i) a fixed total output is produced by the industry and (ii) the total cost of producing the output is fixed. For a simple cost function, whose marginal is linear in logarithm of output, the optimum distribution will be of the Pareto-type.

Key words: Industry performance evaluation; entropy numbers equivalent; optimum output distribution; Pareto distribution.

1. Introduction

It has been often noted that the size distribution of firms, like that of incomes, has a single mode and a highly skewed upper tail.[1] Many authors have regarded the growth of firms as a purely stochastic phenomenon resulting from the cumulative effect of the chance operation of many factors acting independently (see Hannah and Kay, 1977, for further discussions). But the stochastic process theories attribute the observed pattern of the size distribution of firms entirely to the operation of the laws of chance. Thus, these theories rely too little on the economic factors underlying the distribution of firm size.

In this paper we adopt a normative approach to derive the size distribution of firms. Essential to this alternative approach is the Blackorby–Donaldson–Weymark

* For comments and suggestions, I am grateful to J. Graaff, A. Sandmo and participants in a seminar at the University of Karlsruhe. I am also grateful the German Research Foundation for financial help.
[1] We assume that a firm's size is measured by its physical output. The formal discussion applies equally well to any other scalar measure of a firm's performance.

(1982) industry performance evaluation function, a Cobb–Douglas function of total output and an inverse measure of concentration (a numbers equivalent index).[2] Using the Shannon (1948) entropy numbers equivalent index, this paper finds the size distribution of firms which maximises (minimises) the industry performance evaluation function (concentration) subject to the constraints that (i) a fixed total output is produced by the industry, and (ii) the total cost of producing the output is fixed. The second constraint is adopted to hold the resources available to the industry constant, which is appropriate in a partial equilibrium evaluation. Thus we determine the optimal output distribution in a *second best* framework in which a policy maker has direct control over the distribution of output but not over the resources available to the industry.

For a simple cost function, whose marginal is linear in logarithm of output, the optimal distribution has the Pareto density function. Therefore, we have shown how the Pareto distribution, as a size distribution of firms, can be generated from a *second best* optimization framework.

The paper is organized as follows: In Section 2 we discuss the industry performance evaluation function. Section 3 derives the optimum size distribution of firms. We make some concluding remarks in Section 4.

2. The industry performance evaluation function

The number of firms in an industry is indexed by $n \in N$, where N is the set of positive integers. For a given $n \in N$, the set of all output distributions is D^n, with a typical element $\underset{\sim}{x}=(x_1,x_2,\dots,x_n)$, where D^n is the non-negative orthant of the Euclidean n-space R^n with the origin deleted. The set of all possible output distributions is $D:=\bigcup_{n\in N} D^n$. For all $n \in N$, $\underset{\sim}{x} \in D^n$, we write $\bar{x}$ for the sum of the components of $\underset{\sim}{x}$ and $(z_1,z_2,\dots,z_n):=\underset{\sim}{x}/\bar{x}$ for the vector of output shares. For any function $H: D \rightarrow R^1$ the restriction of H on D^n is denoted by H^n.

The Blackorby–Donaldson–Weymark (1982) industry performance evaluation function is defined by $E: D \rightarrow R^1$ where

$$E^n(\underset{\sim}{x}) := (\bar{x})^r (q^n(\underset{\sim}{x}))^{1-r}, \tag{1}$$

$n \in N$, $\underset{\sim}{x} \in D^n$, $0 \le r \le 1$ and $q: D \rightarrow R^1$ is a numbers equivalent index. Thus the evaluation of industry performance has been expressed as a trade-off between total output and concentration.[3] In the limit as r approaches one E^n equals total output; as r approaches zero it is the numbers equivalent.

[2] When all the n firms in the industry have equal market shares a numbers equivalent takes the value n.

[3] This is analogous to the requirement that welfare evaluation of income profiles should involve an explicit statement of the trade-off between efficiency and equity. See Chakravarty (1988).

The Shannon (1948) entropy formula $Q:D\to R^1$ is defined by

$$Q^n(\underline{x}) := -\sum_{i=1}^{n} z_i \log z_i \tag{2}$$

for all $n\in N$, $\underline{x}\in D^n$ with the convention that $0\log 0=0$. Q is an inverse measure of concentration; a highly concentrated industry will take on a lower value for Q and is expected to be closer to the monopoly end of the spectrum from monopoly to competition than an industry with a high value for the index. The Shannon entropy numbers equivalent index is obtained by subjecting the formula in (2) to an exponential transformation. Thus, the two indices are ordinally equivalent.[4]

We now consider analogue to (2) for output distributions defined in the continuum. Let F be the cumulative distribution function [$F(x)$ is the proportion of firms producing output less than or equal to x] on the interval $[\mu, v]$, where $0\le\mu<v\le\infty$. We write $f(\cdot)$ for the density function of the size distribution of firms and m for the mean output. Then it is easy to see that the continuous translation of Q in (2) is given by

$$Q(F) := -\int_{\mu}^{v} \frac{xf(x)}{m}\log\frac{(xf(x))}{m}\,\mathrm{d}x. \tag{3}$$

3. The optimum size distribution of firms

Assuming that all the firms in the industry face the same technology, the feasibility constraints introduced in Section 1 can be written as

$$\int_{\mu}^{v} xf(x)\,\mathrm{d}x = m, \tag{4}$$

$$\int_{\mu}^{v} s(x)f(x)\,\mathrm{d}x = C, \tag{5}$$

where s is the identical cost function for the firms and C is the given cost for producing m amount of output. Clearly, assumptions regarding the technology will determine the shape of the cost function. For example, if the technology is subject to constant returns to scale, then the cost function will have a constant average.

Because output is fixed, maximising the industry performance evaluation function is equivalent to maximising the numbers equivalent or an ordinal transform of it.

[4] Q has been popularised by Theil (1967) as an inverse concentration index. Alternative characterisations of Q or its numbers equivalent form have been carried out by Chakravarty (1988a), Chakravarty and Weymark (1988) and Gehrig (1988).

Definition. An output distribution in the continuum will be called optimum if its density function maximises the objective function (3) subject to the constraints (4) and (5).

Theorem. *The density function of the optimum output distribution is given by*

$$f(x)=\frac{k}{x}\exp[-(am/x)(s(x)-(b/a))], \tag{6}$$

where $k>0$ and, a and b are constants such that $\int_\mu^\nu f(x)\,dx=1$.

Proof. We will invoke the *Euler–Lagrange technique*[5] for proving the theorem. Let

$$L(f)=-\int_\mu^\nu \frac{xf(x)}{m}\log\frac{(xf(x))}{m}\,dx-\lambda_1\left[\int_\mu^\nu xf(x)\,dx-m\right]$$
$$-\lambda_2\left[\int_\mu^\nu s(x)f(x)\,dx-C\right]$$
$$=\log m-\int_\mu^\nu \frac{xf(x)}{m}\log(xf(x))\,dx-\lambda_1\left[\int_\mu^\nu xf(x)\,dx-m\right]$$
$$-\lambda_2\left[\int_\mu^\nu s(x)f(x)\,dx-C\right],$$

where λ_1 and λ_2 are Lagrange's multipliers.

Let $h:[\mu,\nu]\to R^1$ be any function such that $\int_\mu^\nu h(x)\,dx=0$. For any arbitrary $\otimes$ denote $L(f+\otimes h)$ by $g(\otimes)$. If $Q(F)$ attains the maximum for some f, then $g(\otimes)$ attains the maximum for $\otimes=0$. Now

$$g(\otimes)=\log m-\frac{1}{m}\left[\int_\mu^\nu x(f(x)+\otimes h(x))\log(x(f(x)+\otimes h(x)))\right]dx$$
$$-\lambda_1\left[\int_\mu^\nu x(f(x)+\otimes h(x))\,dx-m\right]$$
$$-\lambda_2\left[\int_\mu^\nu s(x)(f(x)+\otimes h(x))\,dx-C\right].$$

$$g'(\otimes)=-\frac{1}{m}\int_\mu^\nu xh(x)\log(x(f(x)+\otimes h(x)))\,dx$$
$$-\frac{1}{m}\int_\mu^\nu x(f(x)+\otimes h(x))\frac{xh(x)}{x(f(x)+\otimes h(x))}\,dx$$
$$-\lambda_1\int_\mu^\nu xh(x)\,dx-\lambda_2\int_\mu^\nu s(x)h(x)\,dx.$$

[5] See Courant and Hilbert (1953) for a discussion.

Since $g'(0)=0$, we have

$$-\frac{1}{m}\int_{\mu}^{\nu} xh(x)\log(xf(x))\,\mathrm{d}x-\int_{\mu}^{\nu}\left(\lambda_1+\frac{1}{m}\right)xh(x)\,\mathrm{d}x$$
$$-\lambda_2\int_{\mu}^{\nu} s(x)h(x)\,\mathrm{d}x=0. \tag{7}$$

We rewrite (7) as

$$\int_{\mu}^{\nu}\left[\frac{x}{m}\log(xf(x))+\left(\lambda_1+\frac{1}{m}\right)x+\lambda_2 s(x)\right]h(x)\,\mathrm{d}x=0. \tag{8}$$

Now (8) holds for all h such that $\int_{\mu}^{\nu} h(x)\,\mathrm{d}x=0$. Therefore, we have

$$\frac{x}{m}\log(xf(x))+\left(\lambda_1+\frac{1}{m}\right)x+\lambda_2 s(x)=\lambda_3, \tag{9}$$

where λ_3 is some constant. From (9) we get

$$f(x)=\frac{1}{x}\exp\left[-m\left(\lambda_1+\frac{1}{m}\right)+\frac{m\lambda_3}{x}-m\lambda_2\frac{s(x)}{x}\right]. \tag{10}$$

Clearly we can rewrite (10) as

$$f(x)=\frac{k}{x}\exp\left[-\frac{am}{x}\left(s(x)-\frac{b}{a}\right)\right], \tag{11}$$

where $k>0$ and a and b are constants such that $\int_{\mu}^{\nu} f(x)\,\mathrm{d}x=1$.

To ensure that f given by (11) maximises $Q(F)$ we need to verify the second order condition $g''(0)<0$. Now

$$g''(\otimes)=-\frac{1}{m}\int_{\mu}^{\nu}\frac{(xh(x))^2}{x(f(x)+\otimes h(x))}\,\mathrm{d}x \tag{12}$$

which shows that

$$g''(0)=-\frac{1}{m}\int_{\mu}^{\nu}\frac{xh^2(x)}{f(x)}\,\mathrm{d}x<0. \tag{13}$$

Thus f given by (11) is associated with a maximum of $Q(F)$. This completes the proof of the theorem. □

It is important to note that exact identification of the parameters appearing in (11) has not been made. The parameters k, a and b are determined by constraints (4) and (5) and the condition $\int_{\mu}^{\nu} f(x)\,\mathrm{d}x=1$ as soon as the cost function $s(x)$ is known.

A second feature is that the general formula in (11) shows the optimality of a long upper tail although we maximise the industry performance evaluation function. That is, the theorem shows that in the optimal situation firms of small as well as of large sizes will exist side by side. To explain this feature let us consider two cost functions s_1 and s_2 where s_1 has a constant average δ and s_2 has a declining average over the interval $[\mu, \nu]$, $0<\mu<\nu<\infty$. Clearly, for all $x>x_0$, where x_0 is the level of output at which $s_1(x)=s_2(x)$, s_2 has a smaller marginal than s_1. We can also show that for all $x>x_0$, $f_1(x)<f_2(x)$, where f_i is the optimal density corresponding to the cost function s_i, $i=1,2$. That is, f_2 has a thicker upper tail than f_1. This means that if efficient production is to be ensured then the optimal situation should allow the existence of more large firms (producing output more than x_0 and controlling a reasonable portion of aggregate output) whenever s_2 is adopted against s_1. But by way of doing this we introduce some inequality into the size distribution of firms. Thus, the optimal situation shows a trade-off between efficiency in production and equity in the size distribution of firms. Equivalently, we say that optimality of a long upper tail is compatible with trade-off between production efficiency and distributional equity.

We may now illustrate the general formula in (11) by an example. Suppose that the cost function is of the form

$$s(x) := \frac{a}{b} + x \log x. \tag{14}$$

The marginal associated with (14) is linear in $\log x$ and takes non-negative values if $x \geq \mathrm{e}^{-1}$. It can therefore be said that the threshold parameter μ for the optimal output distribution corresponding to the cost function in (14) is given by e^{-1}. We also note that for this cost function $(b/a)-\mathrm{e}^{-1}$ can be interpreted as the fixed cost of production. Substituting $s(x)$ given by (14) in (11), we have

$$f(x) = kx^{-r-1}, \quad \mathrm{e}^{-1} \leq x \leq \infty, \tag{15}$$

where $r=am>0$. Using the fact that $\int_{\mathrm{e}^{-1}}^{\infty} f(x)\,\mathrm{d}x = 1$, we can rewrite $f(x)$ in (15) as

$$f(x) = r\,\mathrm{e}^{-r} x^{-r-1}, \quad x \geq \mathrm{e}^{-1} > 0, \tag{16}$$

which is the Pareto density function with the threshold parameter $\mu=\mathrm{e}^{-1} \cdot r$ here becomes the Pareto inequality parameter. It is obvious that the distribution has a finite mean if $r>1$.

4. Concluding remarks

Most of the existing characterisations of the size distributions of firms place too little reliance on economic theory. In this paper we considered the problem of generating the most preferred output distribution in the continuum that maximises the industry performance evaluation function subject to a set of feasibility constraints. A special case of the framework establishes the Pareto distribution as the most preferred size distribution of firms. Thus, we have a characterisation of the Pareto distribution as the size distribution of firms without recourse to stochastic foundations.

References

C. Blackorby, D. Donaldson and J.A. Weymark, A normative approach to industrial performance evaluation and concentration indices, Europ. Econ. Rev. 19 (1982) 89–121.

S.R. Chakravarty, Extended Gini indices of inequality, Int. Econ. Rev. 29 (1988) 147–156.

S.R. Chakravarty, On the separable industry performance evaluation function, Econ. Lett. 27 (1988a) 179–182.

S.R. Chakravarty and J.A. Weymark, Axiomatisations of the entropy numbers equivalent index of industrial concentration, in: W. Eichhorn, ed., Measurement in Economics (Physica-Verlag, Heidelberg, 1988).

R. Courant and D. Hilbert, Methods of Mathematical Physics, Vol. 1 (Inter-Science, New York, 1953).

W. Gehrig, On the Shannon–Theil concentration measure and its characterisations, in: W. Eichhorn, ed., Measurement in Economics (Physica–Verlag, Heidelberg, 1988).

L. Hannah and J.A. Kay, Concentration in Modern Industry: Theory, Measurement and the U.K. Experience (Macmillan, London, 1977).

C.E. Shannon, A mathematical theory of communication, Bell System Technical J. 27 (1948) 379–423; 623–656.

H. Theil, Economics and Information Theory (North-Holland, Amsterdam, 1967).