Aggregative Activity Analysis*

H. S. Houthakker

July 15, 1953

1. The aggregation problem in production theory consists in the formulation of relations between inputs, outputs and prices that apply to a whole industry or economy rather than to a single firm. Investigations of this problem in its most general form (cf. [1] and papers there quoted) have so far led to few positive results; moreover these investigations did not yet show the emphasis on non-negativity which stands out in more recent work on production and allocation [2]. Here an attempt will be made to discuss aggregation with the aid of more specific assumptions derived from activity analysis, and thereby to shed some light on such well-known empirical regularities as the Cobb-Douglas production function and the Pareto distribution.

2. We consider a set of "production cells", which may sometimes be thought of as firms but in other situations as departments of a firm or even as individual laborers or machines. Each cell is capable of producing one or more outputs $x_0, \ldots, x_k$ by means of one or more inputs $x_{k+1}, \ldots, x_m$ (inputs are regarded as negative outputs). Some of these inputs, say $x_{k+1}, \ldots, x_n$ ($n < m$)

* I wish to acknowledge the benefit I derived from discussions with Lester O. Telser and several other colleagues.
are available in variable amounts, whereas the remainder \( x_{n+1}, \ldots, x_m \) is fixed for the cell and period considered. For each cell there is a number (which may be one or infinite) of "production possibilities" describing what combinations of inputs and outputs can be obtained. One of these possibilities will always consist in not producing at all, so that the fixed inputs will yield no outputs, for it is assumed that there will be no output without a positive input of at least one variable and at least one fixed factor.

3. If the production possibilities actually utilized by each cell are known we can sum the outputs and variable inputs involved over all cells and obtain total outputs and inputs \( x_1, \ldots, x_n \) (there is no particular reason for summing the fixed inputs). The main concern of this paper will be with these totals. Clearly not every vector \( X \) is achievable with the given distribution of fixed inputs between cells, and some of those that are possible will be more efficient than others. We are therefore led to extend Koopmans' notion of an efficient point [2] to the production possibilities of a set of cells (an "industry").

Definition. An input - output vector \( X = (x_1, \ldots, x_n) \) for a set of production cells is efficient if, given the distribution of fixed inputs \( x_{n+1}, \ldots, x_m \) between cells, \( X \) is possible and for no possible \( X' \uparrow X \) we have \( X' \geq X \).

Next we give a simple proof of a well-known result:

**Theorem I:** If each cell chooses that production possibility which maximizes its net profit \( \sum_{i=1}^{n} p_i x_i \), then the total input - output vector \( X \) is efficient.

**Proof:** Evidently \( X \) is possible. In order to get another vectors \( X' \) some cells will have to change their inputs and outputs, so that their net revenue will become \( \sum p_i (x_i + \Delta x_i) \leq \sum p_i x_i \), since otherwise they would have adopted
the new possibility previously. Total revenue for all cells will be 
\[ \Sigma p_i (X_i + \Delta X_i) - \Sigma p_i X_i = \Sigma p_i X_i'. \]
Hence it is not possible that \( X_i' \geq X_i \); at least one of the \( X_i' \) must be less than the corresponding \( X_i \).

It will be noticed that this theorem assumes nothing more about the 
production possibilities for each cell than that net revenue reaches its 
maximum. An immediate corollary is that among all efficient vectors the 
one attained gives maximum total profit.

Definition. A production function is a function \( F(X_1, \ldots, X_n) \) which is zero 
if and only if \( (X_1, \ldots, X_n) \) is an efficient input - output vector.

We can therefore say that for the industry as a whole the following 
maximum problem is solved by the price mechanism:

\[ \begin{align*}
\max & \hspace{1cm} \sum p_i X_i \\
\text{subject to} & \hspace{1cm} F(X_1, \ldots, X_n) = 0
\end{align*} \]

4. To make use of (1) we have to put some restrictions on the production 
possibilities. It will be assumed that the cells are small and numerous, that 
production possibilities tend to be different between them, that there are no 
indivisibilities apart from those due to the fixed factors, and that cases in 
which two or more activities are equally profitable to a cell can be neglected. 
The more nearly these assumptions are fulfilled, the closer the aggregate 
efficient points will be together. In the limit they will form a continuous 
surface in n-dimensional space, which as a result of (1) will be convex. By 
making even stronger assumptions (which we are not yet prepared to specify) 
\( F \) can be regarded as a twice differentiable function; (1) then leads to

\[ p_i = \lambda \frac{\partial F}{\partial X_i} = 0 \]
the well-known property of marginal productivity and similarly to

\[ \frac{\partial x_1}{\partial p_1} \neq 0 \]

and

\[ \frac{\partial x_1}{\partial p_j} = \frac{\partial x_1}{\partial p_1} \]

6. The relation between the aggregate production function \( F \) and the production possibilities of the individual cells will first be illustrated by a simple example. Each cell produces one output \( x_1 \) with the aid of one variable output \( x_2 \) and one fixed output \( x_3 \). Each activity requires \( a_2 \) units of \( x_2 \) to produce \( a_1 \) units of \( x_1 \) per unit of \( x_3 \), where \( a_1 \) and \( a_2 \) vary over all activities and cells. The prices of \( x_1 \) and \( x_2 \) are \( p_1 \) and \( p_2 \). See Fig. 1.

The dots represent the production possibilities of a particular cell, which include the possibility of idleness. Two parallel lines are drawn, the one to the left having an equation \( p_1 a_1 + p_2 a_2 = 0 \) and thus indicating zero profit. The other line goes through the activity which gives largest profit, since no dot lies to the right of it. Of course the two lines may coincide, in which case idleness will be the possibility utilized.
Diagrams like Fig. 1 can be drawn for each cell; under our assumptions one point will be found for each. For the 1-th cell this point will be called $a^i_1$; it implies an input $a^i_2 x^i_3 = x^i_2$ and an output of $a^i_1 x^i_3 = x^i_1$. If all these diagrams are superimposed the $a^i$ will be found distributed over the relevant quadrant of the $(a_1, a_2) = \text{plane to the right of the zero-profit line with slope } p_2/p_1$ ($p_1$ can be put equal to one so that the slope becomes $p_2$). The total value of $x_3$ summed over all the cells who have their maximum profit at a point $(a_1^*, a_2^*)$ when prices are $(1, p_2^*)$ will be called $\psi(a_1^*, a_2^*, p_2^*)$. Consequently

\begin{align*}
    x_1 &= \int_0^\infty \int_{-\infty}^0 a_1 \psi(a_1, a_2, p_2) \, da_1 \, da_2 \\
    x_2 &= \int_{-\infty}^\infty \int_0^\infty a_2 \psi(a_1, a_2, p_2) \, da_1 \, da_2
\end{align*}

This is a parametric representation of the aggregate production function $F(x_1, x_2)$.

In the general case, with outputs $x_1, \ldots, x_k$ and variable inputs $x_{k+1}, \ldots, x_n$ we get

\begin{align*}
    x_1 &= \int_0^\infty \cdots \int_0^\infty \int_{-\infty}^0 \cdots \int_{-\infty}^0 a_1 \psi(a_1, \ldots, a_n, p_2, \ldots, p_n) \, da_1 \cdots da_k \, ds_{k+1} \cdots ds_2
\end{align*}

In (6) the input–output coefficients are still expressed per unit of a fixed input, but if there is more than one fixed input this should be interpreted as "per unit of the selected fixed input that can actually be used given the availability of the other fixed inputs".
7. Since (6) gives the aggregate production function in parametric form it is not very transparent. However, in particular cases it is possible to eliminate the parameters and obtain a direct link between the frequency function \( \psi \) and the production function \( F \). Consider the case where there is only one output \( x_0 \) and two variable outputs \( x_1 \) and \( x_2 \) and each cell can use them only in fixed proportions, viz. \( a_1 \) units of \( x_1 \) and \( a_2 \) units of \( x_2 \) for one unit of \( x_0 \) (it is now no longer necessary to standardize on a fixed input); prices are \( p_0, p_1, p_2 \). There are now two production possibilities for each cell, indicated by

\[
\begin{align*}
\psi(a_1, a_2) & \quad \text{if } p_0 + a_1 p_1 + a_2 p_2 \left\{ \begin{array}{c} x \\ \leq \end{array} \right\} 0
\end{align*}
\]

If we integrate over the region of non-negative profit it is not necessary to regard \( \psi \) as depending on \( p_1 \) because the choice between the two possibilities is then already taken into account. For the industry as a whole we then obtain

\[
\begin{align*}
X_0 &= \int_{-p_0}^{p_0} \int_{-p_0}^{p_0} \psi(a_1, a_2) \, da_2 \, da_1 \\
X_1 &= \int_{-p_0}^{p_0} \int_{-p_0}^{p_0} a_1 \psi(a_1, a_2) \, da_2 \, da_1 \\
X_2 &= \int_{-p_0}^{p_0} \int_{-p_0}^{p_0} a_2 \psi(a_1, a_2) \, da_2 \, da_1
\end{align*}
\]
Put \( a_1 = -b_1 \), \( a_2 = -b_2 \), \( P_1/p_o = \pi_1 \), \( P_2/p_o = \pi_2 \), then for instance

\[
-x_1 = \int_0^{1/\pi_1} \int_0^{1 - b_1 \pi_1} b_1 \varphi(b_1, b_2) \, db_2 \, db_1
\]

Let \( \varphi \), the frequency distribution of the input coefficients, now be of the "generalized Pareto" type

\[
\varphi = A b_1 \alpha_1 b_2 \alpha_2 \quad \lambda, \alpha_1, \alpha_2 > 0
\]

then

\[
-x_1 = A \int_0^{1/\pi_1} \int_0^{1 - b_1 \pi_1} b_1 \alpha_1 b_2 \alpha_2 \, db_2 \, db_1
\]

\[
= A \int_0^{1/\pi_1} b_1 \alpha_1 b_2 \left( \frac{1}{\alpha_2 + 1} \right) \frac{1 - b_1 \pi_1}{\pi_2} \, db_1
\]

\[
= A \frac{1}{\alpha_2 + 1} \int_0^{1/\pi_1} b_1 \alpha_1 b_2 \left( \frac{1 - \frac{b_1 \pi_1}{\pi_2}}{\pi_2} \right) \alpha_2 \, db_1
\]

Put \( \pi_1 b_1 = t \) and simplify

\[
-x_1 = \frac{A}{(\alpha_2 + 1) \pi_1 \pi_2} \int_0^{1} t^{\alpha_1+1} (1-t)^{\alpha_2+1} \, dt
\]

\[
= \frac{A}{(\alpha_2 + 1) \pi_1 \pi_2} B(\alpha_1 + 2, \alpha_2 + 2)
\]
Similarly

\[ X_0 = \frac{A}{(\alpha_2 + 1) \prod_1 \alpha_1^{1+1} \prod_2 \alpha_2^{1+1}} B(\alpha_1 + 1, \alpha_2 + 2) \]

\[ -X_2 = \frac{A}{(\alpha_1 + 1) \prod_1 \alpha_1^{1+1} \prod_2 \alpha_2^{1+2}} B(\alpha_1 + 2, \alpha_2 + 2) \]

Since

\[ X_0 = \frac{A(\alpha_1 + \alpha_2 + 3)}{(\alpha_1 + 1)(\alpha_2 + 1) \prod_1 \alpha_1^{1+1} \prod_2 \alpha_2^{1+1}} B(\alpha_1 + 2, \alpha_2 + 2) \]

we obtain

\[ X_0 = C (-X_1) \frac{\alpha_1^{1+1}}{\alpha_1^{1+2} + 3} (-X_2) \frac{\alpha_2^{1+1}}{\alpha_2^{1+2} + 3} \]

where C is a constant. Putting the exponents equal to \( \beta_1 \) and \( \beta_2 \) we see that

\[ X_0 = C (-X_1)^{\beta_1} (-X_2)^{\beta_2} \]

which is a non-homogeneous Cobb-Douglas production function, thus seen to be related to the Pareto distribution.
Footnotes

1/ The definition of a cell will depend on what inputs are regarded as fixed. If, for instance, certain entrepreneurial resources are so regarded the cell is thereby identified as a firm.

2/ The analysis is exactly the same if there are more than two variable outputs.

References


Errata to ECONOMICS NO. 2085

Aggregative Activity Analysis

p. 4 line 8 from top: for "output" read "input"

p. 6 line 5 from top: for "outputs" read "inputs"

formula (7): for \( \Phi(a_1, a_2) \) read \( \Phi^1 \), where \( \Phi^1 \) is the capacity of the \( i \)-th firm

p. 7 line 1 from top: for \( \frac{p_1}{p_0} \) read \( \frac{p_1}{p_0} \)

second line after formula (11) should read

\[
= \Lambda \int_{0}^{1} \frac{1}{\Pi_1} \alpha_1^{+1} \left[ \frac{1}{\alpha_2^{+1}}, b_2 \right] \Pi_1 \frac{1-b_1}{\Pi_2} \, db_1
\]