

June 14, 1948

A Formal Theory of Aggregation

by Kenneth J. Arrow

1. Foreword.

This paper is an attempt to set forth in abstract terms the nature of the problem of aggregation, a problem which, I believe, will be found to arise in many fields other than economics. The reader will observe immediately the resemblance of the present theory to the papers of Marschak and Andrews and of Haavelmo.¹

1. J. Marschak and W.H. Andrews, "Random Simultaneous Equations and the Theory of Production," Econometrica, vol. 12 (1944), pp. 143-205.
T. Haavelmo, "Family Expenditures and the Marginal Propensity to Consume," Econometrica, vol. 15, (1947), pp. 335-341.

2. The Nature of the Problem.

It is a commonplace of the modern operational viewpoint that no question can be regarded as meaningful unless we can construct a test which will determine whether or not any given answer is the true. The literature of index numbers in economics is vast and the bulk of it is dominated by the metaphysical idea of an all-purpose index number somewhat analogous to the "universal solvent" of medieval alchemy, a requirement which, if it can be formulated at all, would exclude any solution of the problem. This is not to deny a great deal of value in the early literature of index numbers, but, as is also often the case in economic theory, the most important contributions arise from a revolt or a subconsciously correct logic against the expressly stated but general principles. As in most scientific issues the correct formulation of the problem is the most essential step in its solution.

Much of the difficulty in posing the aggregation problem arises from confusion between several interrelated types of aggregates. From one point of view, the end of econometric analysis is the determination of policies to maximize the economic welfare, here conceived as some function of the economic behavior of all individuals. (The last statement is really a definition of the concept "individual" for our purposes.) Under this argument, the aggregates to be forecast under varying possible policies are determined by our normative purposes. Strictly speaking, there is only one aggregate, economic welfare, but usually this aggregate is expressed as a function of several other aggregates, and it is usually more convenient to forecast the latter. If we have a complete system (Walras, Hicks) for describing the behavior of individuals, we can eliminate the microvariables with the aid of the definitions of the aggregates to obtain a complete system. (This description blinks the difficult problem of aggregating the predetermined variables; of this, more anon.) This is essentially the viewpoint of May,² and the present essay is a generalization of this approach.

2. K. May, "Technological Change and Aggregation," Econometrica, vol. 15 (1947), pp. 61-62.

There are, however, two aggregations involved in this process. The first aggregation problem is the choice of the aggregates to be incorporated into the system, here dictated by welfare considerations; the second is the formation of the set of parameters of the relations among the first class of aggregates. These parameters are aggregates of the parameters of the behavior of individuals. The distinction must be made more precise.

We must distinguish between the (economic) behavior of an individual at any given time and his response pattern. Each individual at any given

time is faced with a certain number of facts available to him, termed initial conditions;³ these initial conditions restrict the range of

3. This terminology is that of L. Hurwicz, "Theory of the Firm and of Investment," Econometrica, vol. 14 (1946), p. 116.

alternative possible economic policies. From this range of alternatives, the individual chooses the values of several variables, his decision variables.

The principle of selection of any individual is a characteristic of individuals and one which is presumably stable over time; systematic variation over time should be incorporated in the principle itself, by including time or some other relevant variable as an initial condition. In a deterministic system, the individual's decision variables are single-valued functions of his initial conditions; this set of functions defines his principle of selection or response pattern. In a stochastic system, the response pattern can be defined as the conditional joint distribution of the individual's decision variables given his initial conditions. In this case, constancy of the response pattern over time means stability of this conditional distribution.

Those variables which are decision variables at a given time for an individual are the endogenous microvariables. Those variables which are initial conditions for at least one individual but decision variables at the given time for none are the predetermined microvariables. These are assumed to satisfy the usual assumptions as to endogenous and predetermined variables. We thus have a complete system of endogenous microvariables; the response patterns of all individuals, taken together, completely define the joint probability distribution of the endogenous microvariables, given the predetermined microvariables. We will eliminate the latter from consideration for the moment by incorporating them with the response

pattern. Define the current response pattern of an individual at a given time as the conditional distribution of his decision variables, given the decision variables of all other individuals, holding the predetermined microvariables fixed at their current values. The current response pattern is completely determined by the response pattern and the values of the predetermined variables. The current response patterns of all individuals uniquely define the current joint distribution of endogenous microvariables.

We usually identify decision variables of different individuals as being of similar types; e.g. gasoline consumption by individual A is in some sense an activity of the same type as gasoline consumption by individual B (since we are considering simultaneously all economic behavior, the fact that A may want the gasoline for a truck while B wants it for an automobile will be taken care of). I ignore here problems of interpersonal comparison of utilities; some such comparison must be made if we are to have welfare aggregates.

For any one such type of endogenous microvariable, a frequency distribution over the population can be formed if we are given the values of all the endogenous microvariables at any given time. Similarly the joint frequency distribution of all types of endogenous microvariables can be determined by knowing the values of all endogenous microvariables at a given time. This joint frequency distribution will be termed the current behavior distribution.

Assuming that all variables relevant to economic welfare have been included in the endogenous microvariables, all welfare criteria are uniquely defined by the current behavior distribution. Hence, knowledge of the latter is what is really desired to answer any possible welfare question.

In the present essay, the current behavior distribution corresponds to the aggregates to be forecast.

The probability distribution of the current behavior distribution is defined by the joint probability distribution of the endogenous microvariables which is in turn determined by the current response patterns of all individuals. The estimation of the latter in some aggregate sense (defined more precisely in Section 3) will enable us to make forecasts of the current behavior distribution and hence of all welfare criteria. Therefore, it seems that the solution of the second aggregation problem is the essence of a useful solution of the first.

The estimation of current response patterns really involves two distinct problems: (1) the estimation of the response patterns proper; and (2) the aggregation of the predetermined variables. The latter, which we may term the third aggregation problem, does not seem to have received much discussion.

3. Mathematical Formulation.

Let small letters denote elements of spaces corresponding to capital letters spaces. If X and Y are two spaces, X_Y will be the space of all mappings of Y on X . x_Y will be an element of X_Y , i.e., a particular function in Y with values in X ; and x_y will be the value of the function x_Y for the particular point $y \in Y$. $[X]$ will be a probability or frequency measure on X , $[X]$ the space of all possible such measures. The subscript t will denote time. A Greek letter will denote a set in the space with the corresponding Latin capital. Thus $[X]$ (\mathcal{E}) will be the measure of set \mathcal{E} in space X .

I = space of individuals; it includes all individuals who have (or could have?) existed during the relevant period.

B = behavior space; a point in B describes completely the economic behavior of an individual at a given time. It is assumed that the same components can describe the behavior of all individuals; if some kind of economic activity is not engaged in by a particular individual, the corresponding variable can be set equal to zero.

D = space of current response patterns.

L = space of lagged endogenous microvariables, including all relevant lags.

P = space of response patterns.

E = space of exogenous microvariables.

In this notation d_{it} is the schedule describing at time t the current response pattern for each individual, and $[b_I]_t$ is the joint probability distribution of the endogenous microvariables at time t .

Theorem 1. d_{it} determines $[b_I]_t$.

The space I has a natural measure, the counting measure, so that $[i]$ is known.

Given $[b_I]_t$ and the measure on I , the measure on B is determined.

Theorem 2. $[b]_t(\beta) = \int_{B_I} [i] (b_{it} \in \beta) d[b_I]_t$.

From Theorems 1 and 2, it is clear that d_{it} determines $[b]_t$.

However, $[b]_t$ is invariant under a permutation of the labels on I while d_{it} is not. For a given d_{it} , the conditional measure $[d_I | d_{it}]$ obviously contracts to a single point's having measure 1. Hence, analogously to theorem 2, we have

Theorem 3. $[d | d_{it}](\delta) = [i] (d_{it} \in \delta)$

$[d]_t = [d | d_{it}]$ is invariant under permutations of the labels on I .

Clearly, only the frequency distribution of the current response patterns

(and not their assignment by individuals) determines the distribution of the current behavior distributions.

Theorem 4. $[b]_t$ is determined by $[d]_t$.

The possibility of forecasting b_t , and therefore all welfare criteria, rests then upon the possibility of estimating $[d]_t$, the frequency distribution of current response patterns in the population of individuals. This frequency distribution may be termed the aggregate current response pattern.

We have made the assumption that the current response pattern is determined by the predetermined microvariables and the individual response patterns.

Assumption: $d_{it} = d_{p_i} \times l_{it} \times e_{it}$

Here $A \times B$ denotes the Cartesian product of spaces A and B , $a \times b$ an element in $A \times B$. The response pattern p_i differs from individual to individual but is assumed constant over time. l_{it} and e_{it} are assumed to enter into d_{it} for all individuals; this can always be done trivially.

For any functions, if $f(x) = g(x)$ for all x , f and g are the same function. In our notation, this means we can capitalize small letters on both sides of an equality such as the Assumption

$$d_{it} = d_{P_i} \times l_{it} \times e_{it}$$

By theorem 3,

$$[d]_t(\mathcal{D}) = [d | d_{it}] (\mathcal{D}) = [i] (d_{P_i} \times l_{it} \times e_{it} \in \mathcal{D})$$

For a given ℓ_{it}, e_{it} , define $\Pi_t(\mathcal{S})$ as the set in P for which

$$d_p \times \ell_{it} \times e_{it} \in \mathcal{S}.$$

Then

$$[d]_t(\mathcal{S}) = [i]_t(p_i \in \Pi_t(\mathcal{S})) = [p](\Pi_t(\mathcal{S}))$$

Hence, if ℓ_{it} and e_{it} are known, $[d]_t$ would be completely determined for each t by $[p]$, the frequency distribution of response patterns (e.g., the joint frequency distribution over individuals at marginal rates of substitution in both production and consumption and of elasticities of expectation, etc.).

Suppose $[p]$ is known to belong to a class $[P]$. Working backwards, observed values of $[b]_t$ have a distribution determined by $[p]$. So that we can estimate $[p]$. In practice, we have no observations on $[b]_t$ but estimates $[\hat{b}]_t$. The distribution of these estimates depends not only on $[p]$ but also on the method of sampling. If the latter is known, we may again infer from observations $[\hat{b}]_t$ to $[p]$.

Let us further assume that response patterns are impersonal in the sense that the individual reacts to the current and lesser decision variables and exogenous conditions of other individuals only because of their economic status as revealed by the values of their appropriate variables. E.g., a steel producer A may be guided in his decisions by the assets held by B, not because B has the name B, but because B is a consumer of steel. However, these remarks do not apply to the predetermined variables which directly affect the individual, such as his own past decisions. Therefore, the predetermined variables enter the individual's response pattern p_i only through ℓ_{it}, e_{it} and $[\ell \times e]_t$

If we have estimates of $[\hat{l} \times \hat{e}]_t$ and if we have measurements for a sample of individuals for b_{it}, l_{it}, e_{it} , we can continue to estimate $[p]$ essentially as before, when the above assumptions are made. In brief, an estimate $[b \times \hat{l} \times \hat{e}]_t$ suffices to estimate $[p]$.

The most difficult case is that in which the observations on the predetermined variables are not made on the same individuals, i.e., we have estimates $[\hat{b}]$ and $[\hat{l} \times \hat{e}]_t$ but none of the joint frequency distribution. Then, in effect, we have to estimate $[d]_t$, where $[d]_t \in [D]_t$, $[D]_t$ is restricted by some marginal restrictions: $[d]_t = [p \times \hat{l} \times \hat{e}]_t$, where $[p] \in [P]$, a known class, and $[\hat{l} \times \hat{e}]_t$ is estimated. Also, since $l_{it} = (b_{i,t-1}, \dots, b_{i,t-k})$, certain further relations are implied. It may happen in this case that $[p]$ is not identified unless we are willing to make some assumptions concerning the correlation in the space of individuals between p and e_t . Thus, government transfer payments are apt to be negatively correlated with efficiency as a laborer.

4. The Stochastic Element in the Aggregation Problem.

It may seem surprising that the above discussion of the aggregation problem runs so much in probability terms. There are two chance elements: the random element in the behavior of individuals (which may arise either from "intrinsic" randomness or from omitted variables), and the incomplete sampling of the population. Edgeworth evidently had one or both of these elements in mind when referring to the probability nature of the index-number problem.⁴ May's analogy between the aggregation problem in economics

4. F. Y. Edgeworth, "A Defense of Index Numbers," Papers Relating to Political Economy, Vol. I, pp. 356-368, Macmillan, London, 1925.

and the Kinetic theory of gases points in the same direction.⁵ Indeed,

5. K. May, op.cit., p. 59.

it might well be said that the problems are essentially identical.

One can envisage a perfectly deterministic aggregation problem, where each individual is subject to perfectly deterministic laws, and a sampling is complete. In this case, the problem reduces to the determination of $[p]$ from a priori known class $[F]$. The solution of the problem still depends on the nature of $[P]$. This is reminiscent of Klein's solution of the aggregation problem for production, where the form of the aggregates depended upon the a priori known form of the production function.⁶ Although Klein started from a different viewpoint

6. L. Klein, "Macroeconomics and the Theory of Rational Behavior," Econometrica, vol. 14 (1946), pp. 93-108.

as to aggregation, it seems likely that his results can be reinterpreted in light of the present viewpoint. Leontief's recent work on the interrelation of subsets of variables in a continuous function lays the basis for a generalization of Klein's results.⁷

7. W. Leontief, "A Note on the Interrelation of Subsets of Independent Variables of a Continuous Function with Continuous First Derivatives," Bull. Am. Math. Soc., vol. 53 (1947), pp. 343-350. "Introduction to a Theory of the Internal Structure of Functional Relationships," Econometrica vol. 15 (1947), pp. 361-373.

5. The Parametric Case.

Practically speaking, we will be interested in the case where the class $[P]$ can be described by a finite number of parameters. Usually, we will also assume that P is characterized by a finite number of parameters. In that case $[p]$ is the joint distribution of a finite number of variables, and it is known to belong to a family characterized

by a finite number of parameters.

Roughly, the presupposition is that people are not too dissimilar. This is the same thought expressed in Marshall's "representative firm."⁸

8. L. Hurwicz, op.cit., p. 129 n.

It has been objected that we cannot well express, e.g., the transformation function of every entrepreneur, by a finite number of parameters, for if we could so express those functions, the spatial distribution of the parameters would have to be multimodal in order to explain the existence of distinct industries. This objection does not seem well-founded. The economics of specialization are incorporated in the form of the transformation function. Therefore, the profit-maximizing outputs of different goods can vary quite abruptly with small changes in the parameters.

6. The Macrovariable Approach.

We may define an endogenous macrovariable as a parameter (real-valued function) of the current behavior distribution $[b]_t$. The assumption of the macrovariable approach seems to be that $[b]_t$ can be characterized by a finite number of parameters.

Under the present assumptions, even making the parametric assumptions of Section 5, this last statement will not hold. For $[b]_t$ to be characterized by a finite number of parameters, so must b_{It} be. But in view of the random elements entering into b_{It} , it is clear that there is no reason to expect this to be true.

However, two countervailing points must be mentioned. First, a set of endogenous macrovariables, though not exhausting the information in $[b]_t$, still have a distribution depending on $[p]$ and hence can be used in estimating $[p]$, though the estimation is not optimum. If the

sampling is incomplete, it still remains true that the estimates of the endogenous macrovariables will yield information about $[p]$. A question that can be posed is to choose that set of endogenous macrovariables which will minimize the loss of information subject to cost restraints.

A second point is suggested by consideration of a special case. Suppose that d_{it} is the same for all i , and further suppose that b_{jt} ($j \neq i$) does not enter d_{it} as a variable and that disturbances are independent. Then, if the number of individuals in l is large, the spatial distribution $[b]_t$ will approximate the common value of d_{it} and so will be characterizable by a finite number of parameters. It is not clear how this generalizes to systems where there is mutual interaction, but the kinetic theory of gases is an example of a system where random variation in the individual gives rise to certainty in the aggregate, even though there is mutual interaction among the individuals.