

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Income Size Distributions in the United States, Part I

Volume Author/Editor: Conference on Research in Income and Wealth

Volume Publisher: NBER

Volume ISBN: 0-870-14161-9

Volume URL: <http://www.nber.org/books/unkn43-2>

Publication Date: 1943

Chapter Title: The Why And How Of Distributions Of Income By Size

Chapter Author: Conference on Research in Income and Wealth

Chapter URL: <http://www.nber.org/chapters/c5715>

Chapter pages in book: (p. 5 - 33)

CHAPTER I

The Why and How of Distributions of Income by Size

I PURPOSES

THE AMOUNT of income individuals or families receive determines, by and large, their actions and attitudes. It affects spending and other disposition of the income; and thereby the ways people respond to changing economic fortunes. It, among other factors, explains the recipient's other economic and social activities (trading property, voting a given party's ticket, adding x children to the country's population). Non-overt consequences—attitudes, feeling, etc.—frequently discussed under welfare, economic or general, may also be imputed to income.

The purposes for which distributions of income by size may be studied therefore depend upon the phenomena for which income receipts (or differences in them or any other aspect of the size distribution) are a significant antecedent. Such phenomena are numerous and will multiply as we come to know more about the ramifying influences of income getting and spending. Hence the purposes size distributions may serve are equally and unmanageably numerous. Their enumeration is not attempted here. Instead, we first discuss three broad groups of purposes relating to: (A) what recipients do with their income; (B) the influence of income on the recipient's other economic and social activities; (C) non-overt states imputable to income. We treat each group briefly and indicate a fourth suggested by this discussion. Each may demand a different income concept, recipient unit, and period of income cumulation.

A DISPOSITION OF INCOME

At the end of the period for which income is cumulated, many of the recipient's activities can be measured in amounts that add up

to the aggregate income received: direct outlay on consumer goods; investment involving exchange of general means of payment for more specific claims upon the income or assets of an enterprise; savings, i.e., delegating investment to some enterprise; or plain hoarding. One obvious purpose of a size distribution is to trace the flow of income through individuals and families into channels of disposition. Since theirs is the right of disposal, receipts must be traced to them and classified by the amount each receives because, both absolutely and relatively, the channels of disposition are determined largely by the amount received.¹

The disposition of individual and family income must not be confused with the apportionment of a national (or regional) income total among types of use, e.g., among categories of consumers' outlay and of capital investment, which does not require tracing the flow through single households, and consequently does not demand a size distribution of individual and family income. When recipients dispose of income, they deal with enterprises and from the records of the latter we can often gather the amounts going into various forms for the country as a whole or for some region.

To discover how income affects the recipient's use of it demands more: we must study the disposition of income not only for the country or region as a whole, but also for individuals and families or groups at specific income levels. A distribution by size gives us an independent variable in the association to be established between the amount of income and the form of its disposition. Naturally, if such an association is established for *all* individuals and families in the country (or region), one by-product, perhaps among the less important, may be an apportionment of the national or regional income total among types of use.²

Within the broad group of purposes concerned with the disposition of income several subgroups may be distinguished. (a) What dependent variable shall we choose? Shall we define channels of disposition both comprehensively and in detail? We may be interested in all forms of disposition in order to account completely for the flow of income through the unit. Or we may be concerned with the bearing income has upon one type of use; e.g., expenses of

medical care or savings.³ When the type of use is comprehensive enough to be subdivided further (and this is true of what seem to be even narrowly defined uses) our interest may lie in the relation between income and the finest subdivisions, or it may relate to broader categories alone. For example, if we are interested in all forms of income disposition, we may divide them into two broad groups (such as outlay and savings) or set up a large number of subgroups; likewise, if our dependent variable happens to be outlay on safety pins we might conceivably subdivide these pins by size, shape, material, etc.

(b) For what individuals and families do we want both size distributions of income and information on types of disposition? We may include all units in the country or only some; and we may define narrower groups in accordance with various specifications. Most important are the specifications that affect the income concept, the definition of the recipient units themselves, and the period for which income is cumulated. Obviously, size distributions of farm family income, urban family income, professional family income, etc., each designed to establish a relation between income and its disposition, may each call for different income concepts, recipient units, and periods.

(c) Our aim may be to find out what regularities prevail in the relation between size of income and the mode of its disposition as a basis for interpreting changing reality and for reasonable prognosis of the future, but without direct relevance to immediate practical action. Naturally, the relation may not immediately and directly prove persistent; but the hope is that in combination with other information we may be able to demonstrate its relative stability. Or we may study the income-outlay relation in order to affect the disposition of income, either by governmental action (such as taxation) or by other mechanisms (such as advertising or propaganda).

(d) The relation between income receipts and their disposition may itself be of interest as an aid in understanding how uses of income vary or how they may be affected by practical action; or the relation may be merely a first link in establishing the association between income and other consequences comprised under groups B

and c. Obviously, income may affect other activities of recipients largely through its disposition, past or present. For example, a family's criminal record may be due to environmental conditions traceable to spending and saving habits which in turn may be in large degree the consequences of a given income.

These subgroups of purposes are not mutually exclusive but can be combined whenever we study a size distribution with more than one objective. For example, we may wish to establish the relation between the amount of income received and outlay on education for all families; for urban families alone; in order to determine whether there is a persistent relation between income and the educational level of urban families; or to trace the association between income and attitudes on religious freedom. In a similar manner we may combine the purposes under group A in numerous ways.

B OTHER ACTIVITIES AFFECTED

A size distribution may be desired as an independent variable in an association with activities that in themselves do not constitute disposition of income.⁴ While examples have already been given, we may still question whether such an association can be conceived without the intermediate link of the disposition of income. Unless it can, we do not have a separate group of purposes.

The answer lies in the distinction between 'disposability' and current disposition. To illustrate, let us assume that Mr. Smith, receiving an annual income of \$20,000, does not vote for the Socialist nominee for president. The size of his income affects his political activity. But does the influence depend upon what he does with his income? One may argue that Mr. Smith, by attaching importance to the maintenance of his large income, wishes to retain the pleasant consequences of having it at his disposal, and that his desire to retain it led him to oppose a Socialist candidate. Is this income disposition in the sense discussed in the preceding section? Is Mr. Smith's vote determined by his spending three-tenths or two-tenths of his income on food? Six-tenths or one-tenth on a yacht? Should we not differentiate between disposability, i.e., the economic power disposal of income confers, and actual disposition, i.e., the apportionment of the current flow among various uses? Naturally,

income affects economic and social activity only if it is at the disposal of the recipient. This does not mean, however, that its influence is necessarily through a specific apportionment of current receipts, for there may be numerous activities that do not constitute disposition, yet are affected by income in other ways.

It would be next to impossible to classify these activities, i.e., those dependent variables for which size of income is decisive, by any substantive criterion (e.g., political activity, crime record, family habits). All social activities of individuals and families are under the sign of the income level, although for many, size of income is only one and perhaps not the most decisive of several determinants. Yet it is important enough to create an eager demand for size distributions by students concerned less with the disposition of income than with other activities of individuals and families.

But there is a great difference between the meaning attached to income size in these cases in group B and those in group A. In the latter, income is a variable whose very amount is translated into equivalent consequences of interest, regardless whether it symbolizes other economic and social factors. In the former, it is often treated as a symbol of economic status, i.e., a gauge of a whole congeries of economic relations between the given unit and others, pertaining to the unit's relative economic power. In group A the dependent variable is directly related to the current income of individuals and families. In group B it comprises activities less directly related to current income proper and hence likely to be associated with income as a symbol or gauge of a congeries of other variables.

Given a substantively defined group of activities that income is assumed to affect, the general principles that differentiate subgroups under A can, with minor modifications, be applied here. (a) It is not relevant whether we attempt to trace completely income flow into income use, but we may be concerned with how specifically we can spot the effect of income on the dependent variable. (b) We can also classify according to how comprehensively we define the groups of units for which an association between income and a given type of activity is sought. (c) Similarly,

we may classify according to whether the association is to be one of the relatively stable elements for purposes of scientific analysis or a basis for practical action. (d) The relation of income to activities other than its disposition may be studied either for itself or as a link in an association between income and other consequences.

C INCOME AND WELFARE

Could welfare or illfare, satisfaction or dissatisfaction, physical well- or ill-being, degree of psychological integration or conflict be measured, the association between them and income would be similar in type to those in groups A and B. We would again have two variables, of which one—the amount of income received—was studied as a determinant of the second—the state of the recipient. In their association the effect of income may be direct, via its disposition, or via other activities ascribable to it. This granted, however, the types of association and the purposes for which size distributions are desired would be analogous to those in groups A and B.

But it is characteristic of such non-overt states that they are not measurable, at any rate such, usually referred to as economic or general welfare, as are of primary concern to economists and other students of social problems.⁵ We have then associations in which only one variable, income, is given. By refusing to measure the absolute level of the dependent variable, and by postulating standard recipient units whose disposal of income is hypothetical (although naturally the hypotheses may be based upon observation of how recipients actually behave), we are able to ascribe some states to certain income levels or differences in income level.

Studies of economic welfare that utilize data on income size follow divers procedures. Some, disregarding the variety of ways in which recipients may respond to their income receipts and hence the range of satisfaction and dissatisfaction they may derive from the same amount of income, set up a budget for a certain living standard and establish whether the recipient's income would cover the items in it or not. For example, a minimum cost of living is calculated for a standard family for a given period and compared with its income to show whether positive welfare accrues. Other studies go further and for standard recipient units translate income

differentials into welfare differentials, with the qualifications that income differentials themselves reveal the sign rather than size of the welfare differentials. Other studies go even further, and, dropping this qualification, measure inequality of income receipts as inequality of welfare.

The studies under c call, therefore, for different bodies of data on income size and other variables. In some the assumptions allow states of welfare or illfare to be imputed directly to income; in others the connection is to be found solely through an actual or hypothetical disposition of income (not only disposability, which, of course, is assumed in all cases). In some the standard unit and the conditions under which income is translated into the dependent variable are defined specifically, calling for narrow definitions of the income concept, recipient unit, etc. In others a broad and vague association between differentials in income and in welfare, etc., is assumed. In some, the assumption of homogeneity of units and the device of a standard unit are applied to all individuals and families in the country; in others, to narrower groups.

Several principles of differentiation by which subgroups of purposes under A and B are formed cannot be applied here. A connection between income and welfare established through assumptions that cannot be checked does not constitute a relation whose stability under changing conditions can be tested. We have rather an evaluation of the size distribution by strongly normative assumptions; and if these norms are valid, they may provide reasons for practical action intended to modify the size distribution itself. Nor can such a relation between the recipient's non-overt state and income serve as a link to connect income with other, observable, consequences. Since the dependent variable in this relation is imputed, it can be retranslated solely in terms of additional assumptions; it cannot be used as an observational basis by which to associate income with observable consequences at further remove.

D INCOME AS EFFECT

In the groups of purposes discussed under A, B, and C income size is an independent variable, that is, from what we know about it we can learn something about its effect on actions or states of

recipients. But income is itself an effect. A size distribution, eventually of interest because of the effect the income produces, may be of immediate interest as a measure of a dependent variable.

A category comprised of size distributions from which it is desired to measure *effects* of certain recognizable factors may seem superfluous. Why be interested in the effects of factors determining size if not because size itself has far-reaching effects? And if it does, how can an attempt to explain it as a *result* yield different criteria from those yielded by considering income as a *cause*?

The answer is that our goal is a size distribution so prepared as clearly to reflect income as a cause. But many uses of the correlation between income size and its consequences depend in turn upon what we know of the factors determining the size of the income itself; yet the identification of these determinants may call for size distributions different from those called for by attempts to study the bearing of income size upon actions or states of recipients. We might designate the former analytic because they tend to reveal the factors that determined them; in the latter, which we need for groups A, B, and C, the results of determinants have already been synthesized.⁶ We therefore designate them synthetic.

Analytic distributions are desired to explain the origin of synthetic distributions. Hence all subgroups under groups A, B, and C, so far as they imply differences in the characteristics of the synthetic distributions needed, are pertinent also to group D. In addition, we may, for any given synthetic distribution whose origin is to be analyzed, envisage various categories of relevant analytic distributions. These subgroups are based upon the analysis of (a) the income total (to show how the size distribution of the comprehensive total is derived from the size distributions of components; e.g., wages, rent); (b) the recipient unit (to show how the aggregate income of the comprehensive unit, e.g., family, is the sum of the incomes of its members, e.g., individuals); (c) the level at which the synthetic distribution is taken (to show how a distribution of differentials is derived from a distribution of aggregates—the former to serve in turn as the independent variable in further study); (d) the period of cumulation for the synthetic distribution

(to show how total income for the longer period is the sum of receipts for the constituent shorter periods).

II BEARINGS

The purposes outlined prescribe the characteristics of the size distributions: (1) the concept of income that leads us to include and exclude certain types of receipts; (2) the exact definition of the unit that receives the income as thus made up; (3) the particular aspect of income that is to be measured, i.e., whether totals for each recipient, a frequency distribution of recipients by the size of their income, or some derived aspect of that frequency distribution, etc; (4) the period for which income is cumulated.

Naturally, these characteristics may vary only if the concepts are not held to strict definition. If we decide in advance that income signifies one certain type of receipts and no others; that family describes a rigidly defined unit; that what we want is a frequency distribution by size; and that we must deal with annual income, then, regardless of its purpose, a size distribution of family income can have only one meaning. But its relevance and usefulness for the purposes to which it may be applied will vary. We proceed here on the assumption that, while dealing in general with a size distribution of family income, we have a range of choice within which to define income and family, and to select the period of cumulation and the aspect of income to measure.

A THE INCOME CONCEPT

Whatever connotations the concept of the income of a family (or any recipient unit) may carry, their common and central core may be described as the net accretion of separable means of command over goods. It is net in that it excludes flows that offset the recipient's loss of goods incurred in the earning of income. It relates to means separable from the recipient proper in that an increase in command due to some improvement of individual capacities, having no objective and measurable form, is excluded.⁷

This common element in the definition of income accepted, such net accretion to command over goods may or may not be further qualified in many ways. The range is from the narrower defini-

tions, in which definite types of net accretion are alone included (e.g., only those that result from participation by the recipient in productive activities, productive narrowly defined; or only those that are relatively regular and forecastable; or only those that take the form of money; or only those that come from enterprises), to the widest, in which all types of net accretion are included. Sometimes even the concept of netness may be widened and income made to include accretions that may be offset by certain changes in the individual's wealth position (these changes being disregarded for the short period studied).

Although using the term 'income' to designate different concepts creates confusion, we can readily see why it is done. If, for example, we wish to establish relatively stable relations between income and, let us say, expenditures on staple foods (to serve possibly as a basis for forecasting their amount), we should perhaps confine income to service earnings and exclude not only such items as capital gains but even some property income items such as dividends. If we seek to foresee short term changes in expenditures on medical care, which for a given family are intermittent and may call for emergency mobilization of all its economic resources, we may deem it advisable to include under family income not only all service and property income receipts, capital gains, etc., but even amounts borrowed or proceeds from property liquidation during a given brief period.⁸ Similarly, the desire to establish the bearing the amount of income received has upon its disposition or other activities as an immediate basis for practical action may dictate another income concept;⁹ and likewise when one wishes to translate income size into states of the recipients or to single out factors producing a given type of size distribution.

True, one can usually no more than surmise what income concept is most relevant to and potentially useful for the purpose at hand. And until we have information on both the amount of the income and the dependent variable and have completed the analysis we possess no adequate basis upon which to infer that this and not another concept is more suitable. Caught in this vicious circle, it may be expedient to define income comprehensively, demanding at the same time that all significant income subcomponents be dif-

ferentiated; were this done, one could experiment with income concepts, choose those that yield results most relevant to the given purposes, and make estimates with which other data can be combined most easily.

As a matter of fact, this is the way our knowledge progresses: mustering results from numerous studies, we progressively refine our definitions of variables that are determinants. But to demand that each study of the distribution of income by size be comprehensive, embracing a variety of income concepts, is to preach a counsel of perfection. A comprehensive concept of income is not one for which information is readily obtained: on the contrary, the measurement and recording of its marginal items is laborious (e.g., net income from certain property items, capital gains and losses, especially when not realized). And, of course, detailed recording of subcomponents is a heavy additional burden on any study designed to collect primary income data. For these reasons studies collecting data on income by size naturally choose, on the basis of tentative and conjectural knowledge of the relation between income and the dependent variables they are interested in, the income concept most suitable for their purpose (modifying it in the light of practical exigencies). The great variety of such variables and of purposes the associations may serve means a multiplicity of income concepts, different from one type of study to the next.

Nor, by making it strictly equivalent to direct shares in the national income total, can we call for a uniform income concept in all and sundry distributions by size. The goal of most studies in which a size distribution is desired is not to show what shares of a given national (or regional) total flow directly from enterprises to individuals and families in the country (or region); it is rather to measure such flow to these units as represents net accretion to economic power (relevant to the dependent variable at hand), regardless whether this accretion is due to such participation in the activities of economic enterprises as warrants direct receipt of shares of national income (or aggregate income payments). True, a size distribution of national income (or aggregate payments) is an important analytic distribution: it shows how participation by individuals and families in the productive activities of economic

enterprises causes differences in the shares flowing directly to them. But these shares may well be redistributed through property exchanges, gambling, taxation, etc; and it is the eventual results of such further redistribution that one may need.

Hence, the income concepts needed in synthetic distributions of individual and family income by size are as a rule not the same as that required for national income. Consequently, the sum of the incomes recorded for the separate units cannot equal national income. It may be smaller because it excludes undistributed profits and losses of enterprises; it may be smaller even than aggregate payments because it may fail to record flows credited to individuals through such associations as insurance companies and savings banks; furthermore, it may omit income from enterprises that is sometimes excluded because of its irregular bearing upon the dependent variable (e.g., certain types of property income). But it is more likely to be larger than either national income or aggregate payments, for it may include: (a) proceeds from productive activities that we should include in national income but usually do not because we cannot estimate them (income from boarders, roomers, and similar intra-family activities); (b) positive counterparts of unproductive activities, but not their negative counterparts (e.g., gains from gambling but not losses); (c) capital gains; (d) property transfers that are not offset (e.g., proceeds from insurance, inheritance, and sometimes direct borrowing from enterprises).¹⁰

In contrast to such a uniquely determined total as national income, many of the totals formed by adding incomes of individuals and families may not have an unambiguous and homogeneous meaning. In other words, although some components of the totals derived from studies of the distribution and disposition of income by size do have a direct relevance and meaning as components of a national income total, others may not. For example, data on consumers' outlay derived from a size distribution of income and based upon a definition of outlay that recognizes only results of such productive activities as are also included in national income are properly a component. But an item such as savings, derived from a size distribution of income that includes capital gains and

positive returns (without offsets) from unproductive activities, is not properly a component of national income.

We come to the same conclusion as in Chapter 4 which suggests a minimum, common denominator concept, but finds it impossible to urge a single and fairly comprehensive definition as adapted to any and every group of purposes. Hence we can mention merely the most obvious essentials of an income concept suitable for various types of size distribution. (a) The diversity of definitions and the multiplicity of items that can be included and excluded counsels the utmost explicitness in formulating the concept and enumerating the items included or excluded. (b) Our meager knowledge of the exact relation between income and any dependent variable suggests the great potential usefulness of defining income comprehensively and segregating numerous subcomponents. (c) And since national income totals have already been studied in considerable detail, with results that may contribute much to the understanding of how a given size distribution arises, we propose that the components of an income concept used in size distributions be at least so detailed that a concept similar to that underlying national income (or aggregate payments) totals can be approximated.

B THE RECIPIENT UNIT AND THE ASPECT MEASURED

In a relation of any two variables, A and B, we must specify the unit for which a given magnitude of A will be recognized as co-existing or co-occurring with a given magnitude of B. This unit is a vital feature of any association: obviously the relation between the length and weight of a leg is different from that between the length and weight of an arm, or between the height and weight of a human body. In general, the unit of identification is so chosen as to constitute a system within which relations between the variables studied tend to be determinate, a system in that it definitely circumscribes the degree to which the relation between the two (and perhaps among other) variables can assume different forms in time or in space. Thus we associate the height and weight of a person rather than the height of a husband and the weight of his wife's favorite dog. In the first association each person can be

treated as a complete system within which possible interrelations of height and weight are definitely circumscribed; in the second, each husband and his wife's dog may well form an open set, with no usefully narrow limits upon the relation that height and weight may assume.

This consideration applies to the large and diversified group of associations in which size of income is one of the two variables. We may wish to deal with the income of each individual because it is individuals who participate in the income producing processes of economic enterprises, receive most types of income, dispose of these income receipts, act in other ways under the influence of income receipts, and experience states of welfare or illfare, satisfaction and dissatisfaction. Each individual can thus be treated as a system, a closed set within which the relations of its income to the consequences or antecedents are definitely circumscribed. Or, since some types of income flow to families rather than to individuals, incomes earned individually by the family's members are pooled for purposes of disposition, and consequences, either overt or imputable, of the amount of income received are more definitely circumscribed when traced through such family units, we may wish to deal with the income that units we designate 'family' (whether each includes one or several individuals) receive. Or, we may go further and create more compound units, if they seem more advantageous for establishing the relation between the amount of income received and such of its antecedents or consequences as are of primary interest.

Here we discuss only size distributions of income among such ultimate units as individuals, families, and other associations of individuals, units that do not constitute economic enterprises. For distributions of this type the units must possess at least three characteristics: (a) they must be clearly identifiable in empirical observation and practical policy; (b) we must be able to trace the income flow to them, i.e., we must be able to associate with each a specific amount of income received; and they must be the units concerned with its ultimate disposition (i.e., disposition outside of economic enterprises) and carriers of overt or imputable consequences; (c) they must be relatively homogeneous so that the

frequency distribution of income by size possesses some meaning in terms of the associations in which size appears as a variable.

The first requirement is easiest met when we choose individuals as the unit of count; no problem of definition or difficulty of observation arises. But the second, and the most compelling requirement, calls for the family rather than the individual as the unit of count, especially if applied, as it must be, to distributions that attempt to account fully for income in a given area. Individuals are a much less suitable unit than families because: (a) some types of income flow to family units, not to individuals as such; (b) to many individuals an exact amount cannot be allocated (e.g., all dependents); (c) income is disposed of largely by the family. On the other hand, a unit larger than the family is just as unsuitable because there is no complete pooling within it of either income receipt or disposition. For the type of distribution under discussion here and in line with the second requirement, we therefore deal with family units alone.¹¹

But the meaning of family as a recipient unit may differ as size distributions of income are made for one purpose or another, particularly in the extent to which it includes individuals who receive their income directly. Should the family Jones include a grandfather who has his own income and who may live with the family or run his own household? Offhand the answer would be to include him if his income is pooled with those of other family members and if his expenses are paid from the common pool. But what degree of pooling is requisite? Some expenses (of the ordinary variety) may not be pooled and the two households may be run separately. On the other hand, when grandfather Jones or his little granddaughter is sick and medical bills pile up, or when a new family car is bought, all may chip in. Thus if we wish to study the bearing of income upon ordinary expenses, grandfather Jones should be excluded from the family unit, but if we wish to study its bearing upon extraordinary expenses, he should be included.

The inclusiveness of the family unit may vary similarly when other activities or non-overt states are the dependent variables. In tracing the association between income and the political activities of Mr. Smith, should the family unit include a wealthy grand-

uncle who may from time to time contribute to Mr. Smith's means of command over goods? ¹² In assigning welfare equivalents to income receipts, should we assume that the community of feeling a family unit has extends to a poor relation living with the household or to a rich relation living outside?

Some limit to the comprehensiveness of the family unit is set by the first requirement, ease of identification and observation, and the third, still to be discussed, homogeneity. Within these limits, however, the range is wide. We are as much at a loss how to define the family unit as the income concept. And the expedient of using the most inclusive is even less practicable; for a family unit most comprehensively defined is hard to identify in empirical work, and any intrusion upon family life in order to discover relations between income receipts and their manifold consequences is vehemently resisted.

A rigidly and uniquely defined family unit can no more be taken over from social studies in which the family is the unit of observation than can an individual income concept be taken over from national income studies. Perhaps sociology provides a single acceptable definition of a family. Even if it does, we may concede the advantage of ascertaining how the income receipts of family units, so defined, differ and what the consequences are, but we must allow in turn that the family be defined with direct relevance to the purposes of the estimates. If the concept of an 'economic' family, in which economic implies the disposition of income, differs from that of a 'social' family, the latter seems scarcely to have a place in a study that deals solely with economic relations. Of course, if there is a great deal of knowledge about the 'social' family that may reveal relations of moment in a given study, this definition should be used, or at least discussed; but it cannot be put forward as the one concept most appropriate for all purposes.

The analogy with income concepts is evident, as is the folly of calling any one definition 'correct'. Before stating the effects various ways of defining the family unit have, we consider the third requirement, relative homogeneity.

So long as data on income are for each recipient unit and we do not try to add them or otherwise transcend the boundaries of

each, whether the unit is homogeneous does not matter. One would study the income of and its effects on the Jones family; similarly for the Smith family, and so on, without assembling the results for the several families. But the moment we combine or generalize, each family must be weighed, either explicitly or implicitly, and the homogeneity of the unit of count or equivalence of the several families questioned.

Since usually we study incomes on a mass basis the variable can practicably be treated only by means of a frequency distribution of the recipients by the size of their incomes (or by means of some derived aspect of that distribution). We must therefore consider what bearing the implicit assumption of equivalence has upon the definition of the family unit.¹³ 'Equivalence' refers, of course, solely to the effects of the variable in the frequency distribution. If, for example, we wish to establish the distribution of families by the size of their incomes as a basis for estimating their distribution by the size of their expenditures on shoes, it may matter little that some families are red haired and others black haired, i.e., that the unit of count is heterogeneous with respect to color of hair; but it will matter that some families have two and others ten pairs of feet to be shod.

Second, the purposes for which an association between the variables (e.g., income and outlay on shoes) is to be established may be relevant in judging whether the unit is homogeneous. For example, family units may not be homogeneous with respect to number of members and geographic location, both of which affect the relation between income and outlay on shoes. This heterogeneity is decisive if the relation is to be used to establish the elasticity of the demand curve for shoes from a cross-section analysis; but may not be if we wish to know how *changes* in the distribution of income from one year to the next affect the proportion of total income spent in the country on shoes. For the latter purpose, heterogeneity may be assumed to remain constant from one year to the next, or at least shift relatively little, so that changes in the income distribution would not be affected by changes in the relative weight of large and small families or of families living in tropical and temperate climates.

Third, the unit of count, since it is empirical, cannot be absolutely homogeneous. No matter how narrowly one may define a family (e.g., confining it to two members of given sex and age, race, residence), it will still be heterogeneous with respect to any bearing that income has. Hence, in practice we strive merely for relative homogeneity: the heterogeneity of the unit should not be so great that its influence upon the effects of income is sufficient to cancel or hide them. For example, a definition of a family generally applicable whether it has one or ten members, lives in Florida or Maine, on a farm or in the city, may still yield a relatively homogeneous unit for a study intended to show the bearing the size of its income has upon how much it spends and saves. The various respects in which the families are not equivalent still do not obliterate or reduce the underlying functional relations between income (as reflected in a frequency distribution) and its apportionment between expenditures and savings—for two reasons: the range of differences in income is far wider than in the elements of heterogeneity; and in the frequency distribution the extreme effects of heterogeneity are lessened by combining, within one and the same income size group, families differing in the number of their members, geographic location, etc.

Yet this requirement of relative homogeneity cannot be so easily dismissed; it bears on both the definition of the family unit and the aspect of income to be measured. On the family unit of count, the first effect, already suggested, is to limit the degree of comprehensiveness with which it will be defined. One obvious reason for not defining 'family' comprehensively (e.g., including all collateral branches, regardless whether they are closely related or live under one roof) is that the range of variation of factors relevant to causes or effects of income receipts becomes so much wider; and each family in itself presents a much greater mixture of variables than if it were more narrowly defined. Another effect is the implication that the family should be so defined that any elements of heterogeneity potentially relevant to the cause or effect of income receipts under study can easily be taken into account.

For example, if residence in the north or south, in the city or on a farm, is a source of considerable heterogeneity, the family should

be so defined that this particular attribute can be segregated; and thus, by adding information on it, one could study the causes or effects of income receipts, assuming that it does not vary. In some cases this additional information may be used without direct conversion of the original unit of count, by bringing it in as an additional variable or set of variables for which the unit that identifies coexistence with other variables is the 'family' as originally defined. In other cases, this additional information may be used to translate the original unit of count into some derived unit, let us say, consumer-unit equivalents, gainfully occupied equivalents, or some compound and derived (rather than directly observable) unit, from which recognizable variables that were a potential source of significant heterogeneity have been removed. Whatever we do, we should so define the immediately observable unit of count that we have a hook large enough to hold whatever variables may influence the meaning of income as cause or effect.

When it is neither desirable to define the family unit too narrowly nor possible to obtain or conjecture information on variables that constitute sources of substantial heterogeneity, we can attempt to limit heterogeneity by our choice of the aspect to be measured. As already suggested, a frequency distribution of families by the size of their incomes may not only summarize information concerning a large number of units but also weaken effects of other variables that are imperfectly correlated with income (or perhaps not correlated at all). Similarly, measures at further remove from the original data for the specific unit of observation, such as standard deviations and coefficients of variation, may weaken even further the possible disturbance by variables that are sources of heterogeneity in the relation of income to its antecedents or consequences. These and other choices of the aspect of income to be measured may curtail heterogeneity, avoiding a change in the definition of the family unit or in the area or groups for which the size distribution is studied.

But they are makeshifts. If, in order to minimize heterogeneity, thereby revealing more clearly the antecedents or consequences of income, we choose to measure an aspect at some remove from the original data, we must make some assumptions concerning the

behavior of the variables that constitute the source of heterogeneity. For example, if with reference to certain consequences we study differences in income over time rather than at a given moment, we may expect more homogeneity only if the sources of the differences are relatively stable over time. But on such matters we have at best merely tentative knowledge; and the feasibility of making plausible assumptions of this type does not help much in meeting the requirement of homogeneity. Unless the specific aspect of income measured is dictated by the characteristics of the dependent variable income is to bring to light, it is only a makeshift for using a family unit for which additional information can be collected on any source of significant heterogeneity and for dividing the frequency distribution into significantly different groups of families.

To summarize: even if we confine ourselves to units for which we can trace the flow of income via the recipient to its ultimate disposition (i.e., disposition outside of economic enterprises) and with whose aid we can account exhaustively for income flow in a given area, the 'family' may still be more or less inclusive within a range limited by the requirement that the unit be easily identifiable; by the need for relative homogeneity if the income information is to be summarized in a frequency distribution (as it must be for a large number of observations); and by the desirability of so defining the unit that other variables, which are potential sources of heterogeneity, can be assigned to it.

The variety of purposes may determine also the aspect of income that will be measured. In some studies, totals are indispensable in that they provide the basis for the relation to be studied or practical action to be undertaken; the purpose of others is better served by some derived indexes, such as relative differences or an average of them. For many purposes, totals for each specific item are indispensable; for others, a frequency distribution by size suffices. Or the aspect to be measured may be selected because of a desire to make the unit of count more homogeneous. Whatever the reason for our choice, if we measure an aspect at some remove from the original data, the homogeneity of the unit of count is affected; which is why these two characteristics of size distributions of income have been discussed together.

Thus we come to conclusions similar to those concerning the income concept. Within the limits set by the requirement of homogeneity and the practical difficulties of observation it is best to define the family unit most inclusively. For it must be the basis of identification not only of income and the effects (or causes) that are of particular interest in the given study, but also, if possible, a unit for which information on other variables may be gathered, variables that may qualify or affect the meaning of income in its diverse bearings. This means that, if possible, the family unit should be equivalent to the 'family' for which considerable information has already been accumulated. Finally, the definition should be explicit concerning boundaries, which is even more difficult than to delimit the income concept.

Conclusions concerning the aspect of income to be measured are somewhat different. The closer the information is to the basic totals, the more it reveals the variety of forces operating. Any abridging of the primary data that reduces this variety, whether intended to make the unit of count more homogeneous or dictated by the specific level at which the dependent variable is to be interpreted, means a loss of information potentially valuable for other purposes and other variables, and is to be viewed as a sacrifice that may not be fully compensated for by the gain for the specific purpose at hand.

C THE PERIOD OF INCOME CUMULATION

To obtain totals as a basis for distributing families by the size of their incomes, receipts must be cumulated for some finite period. Its length is an important feature of the size distribution. For example, if income is cumulated for a long period, the transient factors that may make for differences in income among families will be weakened and the differentials are likely to be smaller than for a short period. Also internal shifting within the distribution, i.e., changes in the relative standing of families, will tend to be less and different. The period of income cumulation must be carefully chosen with a view to the purposes size distributions are to serve.

It may seem at first that the choice is predetermined by the

period for which the antecedents or consequences of income are to be studied. If we wish to ascertain the effect of income receipts upon annual outlay for various types of goods or upon annual savings, it may seem reasonable to estimate income annually. If we wish to influence in some way the disposition of income or measure its other consequences (political actions, etc.) within annual and monthly periods, it may seem reasonable to cumulate income for the same periods.

But this suggestion is invalid. The outlay of a family unit and its political and social actions during a given year (month, week, etc.) or its welfare within a brief time span may well be affected by its income for a much longer interval. The composition, absolute and relative, of a family's budget and its other activities are affected by receipts not only in the given year but also in preceding years and perhaps also by those expected in the immediate future. To establish clearly the consequences of income it may be necessary to cumulate receipts for a period much longer than the one for which the consequences themselves are to be studied.

This statement becomes more self-evident when the erratic and temporally variable character of some income flows is considered. For the large group of independent entrepreneurs, family income may vary widely from one year to the next; yet for obvious reasons, amounts spent in any given year upon goods of various types, the social activities pursued, or even the welfare imputed are likely to vary much less from year to year; hence they bear an irregular relation to the income for any given year. Can one seriously contend that an income of only \$800 in a given year properly measures the effect of income upon the outlay, political views, or welfare of an entrepreneurial family that has been accustomed to an annual average income of some \$5,000 for two decades? Or that the effects of this \$800 income are similar to those on a family that has been accustomed to an \$800 annual income? Certainly, for both families their respective incomes for a longer period may be more relevant: the entrepreneurial family could not maintain its usual standard of living on \$800, while the other could.¹⁴

If, then, family income as a determinant of monthly or annual consequences calls for totals not for equally short but for much

longer periods, the period of cumulation may well differ with dependent variables and with the purpose the relations to be established are to serve. Outlays on certain types of goods respond slowly or not at all to short term changes in income; outlays on others may respond promptly. The means we may wish to use to affect the disposition of income in the way of practical action may or may not respond to short term variations in income. Some purely social activities may reflect the deviation of annual income from income status; others will reflect the latter alone. And if we do not draw similar distinctions among non-overt states it is because our knowledge of them or our imagination does not reach to the point of segregating responses to short and long term levels of income, although we can and do detect differences in response to short term changes between recipients of temporally fluctuating and of temporally stable incomes.

The more comprehensive the concept of income the greater the need for cumulating receipts for a longer period. Since the borderline components (net profits and losses of entrepreneurs, gains and losses on capital assets, gifts, bequests, borrowings) tend to vary, a comprehensive total for a year is likely to be less representative of longer term status than a narrower total. On the other hand, the more inclusive the definition of the family unit the more, other conditions being equal, is the income for a year likely to represent longer term status; for the more inclusive the family the more stable its composition over time and the more likely temporal variations in income from one member or source will offset those in others.

Although recognizing the need for estimating income for periods varying in length, each perhaps longer than the customary period of a month or a year, we may claim that cumulation must be for short periods and the longer period totals compounded of monthly or annual subtotals. The reasons are obvious. First, even if in establishing the effects the family's income has upon its economic and other activities or non-overt states we need to consider its receipts for a fairly substantial period, more than a single total for several years is required. The pattern of income during preceding years itself is important, for the family bases expectations as to future

income on it; moreover, the family's reactions to a certain income for a given year can be analyzed only if it is treated as a deviation from an average level pattern, which in turn can be observed only in subtotals for short periods.

Second, even though a family's expenditures, actions, or attitudes in a given year may be affected more by its income for the preceding decade than for the current year, still the latter is a variable to which specific effects may be attributed; indeed, for certain purposes, they may be far more influential. For example, if we intend to explain *changes* from year to year in outlays on various types of goods, or in various types of social and political action, or in welfare, all in terms of income receipts, *changes* in annual income from year to year may well be the basic independent variable. The necessity of interpreting them in terms of the longer period flow merely means that totals for both long and relatively short periods must be estimated.

Third, and perhaps most important, size distributions of income cannot be explained unless totals are cumulated for fairly short periods. The family's income status is itself subject to longer term, secular movements; income for any given year is, in addition, affected by cyclical and irregular factors. If we are to understand how a size distribution among families is formed, we must be able to measure the changes over time in the family income, sort out both the secular trends (associating them with the secular movements of the industries and enterprises in which the income earners or their property are employed and with the changes in the wealth and productivity status of the individuals themselves) and the cyclical and irregular elements (associating them either with the industry and the enterprise or with conditions peculiar to the family proper).

Considerable information on industries and enterprises as sources of income is available, information in which the distinction between transient and more persistent movements is based on continuous measures of rates of activity for short periods. Obviously the changes in family income that explain differences in the amounts cumulated for given periods can be explained and associated with the factors working through the enterprises and indus-

tries only if data on income also are available continuously by fairly short intervals. Hence whatever the period for which income as a determinant must be cumulated, all purposes would profit from estimates for relatively short periods within the longer. We may, therefore, reasonably require that any study that calls for income information cumulated for a period longer than the shortest feasible (say a year) also call for income for at least one, or better, several, of the shorter periods within the longer (for each unit or as totals or averages). Granted the practical difficulties, in this fashion alone will information basic to the interpretation of income as the dependent or independent variable be obtained. For we need for successive years information on each family's income as well as distributions of families by size of income; and family income can be ascertained only by direct inquiry of the recipient family itself.

III CONCLUDING COMMENTS

This discussion of what bearing purpose has upon the essential characteristics of size distributions of income is necessarily sketchy. The field of our study has been so little cultivated that we have no large stock of substantive information from which to set up clear-cut criteria for judging precisely what income concepts, family unit, and period of cumulation are most suitable to this or that purpose. All we have attempted are tentative suggestions of how purpose impinges upon the characteristics of size distributions of the generic type discussed.

We did not refer explicitly either to the ease or difficulty of getting accurate data on family income, or the supplementary data requisite if we are to find out what determines the size of family income and its consequences. Differences in the ease of obtaining *information on paid and imputed income, on income receipts from socially approved and disapproved activities, for families that can and cannot readily be defined in terms of living under one roof, for recent and customarily used time units, such as a year, and for other periods* are all fairly obvious. Information on specific income components, specific elements in the definition of the family unit, the aspect to be measured, and for deciding upon the period of

cumulation may be harder to get. We have tried to suggest why we want data on the size of family income and the variety of its determinants rather than to indicate what we can get easily. In short, we aim to suggest the desirable, not the available.

Nor was it feasible to discuss in a substantive fashion the data, other than on size of family income, needed to analyze the latter adequately in terms of its antecedents or consequences. That a large variety is involved is obvious; the most we could do was to group purposes for which size of family income is needed and discuss the variant characteristics of the distributions by size. To go beyond these scanty classifications would lead us to analyze investigations into the causes and effects of size distributions of family income, list the variables, and evaluate the results in terms of other variables not taken into account but apparently relevant.

The inclusiveness with which we define the income concept and the family unit, and our choice of the aspect to be measured and of the period of cumulation determine whether some variables are classified under size distributions of family income or under 'other information'. For example, if we exclude capital gains from the income concept and treat them as changes in property values, then in measuring the effect size of income has upon its disposition and other family activities we must consider capital gains under 'other information'. If we exclude from the family individuals who share common household facilities but receive their own income, then we must separate families with and without such members in order to understand the influence size of income has on the activities of the family unit so defined; and this distinction is a variable under 'other information'. If we confine 'size of income' to the flow for short periods and designate totals for longer periods 'income status', then income status is also a variable under 'other information'.

Thus the more narrowly we define the congeries of concepts designated 'size of family income' the more numerous the variables of potential influence on the size of family income or its consequences that must be classified under 'other information'; the more comprehensively we define the term the fewer the pertinent vari-

ables under 'other information'. Our choice of the more comprehensive definition made for the multiplicity of variants under income concept, family unit, aspect to be measured, and period of income cumulation.

Our preference was motivated by the paucity of information and by the futility of insisting upon too narrow and rigid definitions. If we define income, family unit, etc. comprehensively, many more factors determining how income getting and spending affect or are affected by activities and attitudes can be studied. But for reasons often repeated, we do not urge a comprehensive set of concepts for every study regardless of its purpose.

That our preference is wise and justified by our ignorance of size distributions of family income may well be contested. When we try to think through the ramifications of a circumscribed area of study that necessarily deals with a part of a greater whole, we have to weigh the advantages and disadvantages of narrower and wider boundaries. In a field where our stock of knowledge is small and empirical work in its initial stages, the balance seems to favor comprehensive and varied treatment, in the expectation that as we add to and become acquainted with our data, we can narrow definitions and make nomenclature more specific.

N O T E S

¹ This is the aim of what is at present the most widely accepted distribution of income by size among families in this country, viz., that prepared by the National Resources Committee on the basis of data from the Consumer Purchases Study (see Part II, Ch. 11 and 12).

² Even this does not necessarily follow, since the income concept used in the distribution by size may differ substantially from the concept of national or regional income; see Sec. II A.

³ See, e.g., the uses of size distributions of income in the Studies of the Committee on the Costs of Medical Care (Louis S. Reed, *Ability to Pay for Medical Care*, University of Chicago Press, 1933) or in the National Health Survey (see Part II, Ch. 13).

⁴ Among such uses perhaps the most common are associations of income with family fertility rates, with size of family, with incidence of illness and criminality, and with party affiliation and political opinion.

⁵ It is not impossible that study of nutrition, housing, etc., and their effect upon physical and psychological well-being may eventually give us measures of states

of individuals and families as quantitatively precise and specific as the measures of observable actions.

⁶ Examples of analytic distributions are abundantly provided by studies of income tax returns, both federal and state. Yielding distributions by size of wages and salaries, business income, interest, dividends, etc., each taken separately, they show how the distributions of *total* income by size arise (see Part II, Ch. 1-3).

⁷ Intricate problems arise when these concepts of netness and separability are applied to specific items; and there is close interrelation at these points between the concepts of income and wealth. It would lead us far beyond the proper scope of this chapter to discuss these questions in detail although some are inevitably touched upon below.

⁸ See, e.g., the inclusion of proceeds of borrowing in income concepts used in recent health and medical care studies (Part II, Ch. 13) and the California study summarized by Paul A. Dodd and E. F. Penrose in *Economic Aspects of Medical Services* (Graphic Arts Press, 1939).

⁹ See, e.g., the peculiarities of income definition in distributions of tax return data (Part II, Ch. 1-3).

¹⁰ We could force the sum of incomes of all units in a country to equal at least countrywide aggregate payments by including under the latter items such as (a) above, i.e., proceeds from activities that should have been included and were not; and by deducting from the former offsets for items such as (b), (c), and (d), i.e., corresponding losses from incomes of units and enterprises that made these gains possible.

For example, if we include gambling gains in the incomes of winners, we should deduct losses from the incomes of losers. If capital gains are included, 'opportunity' capital losses should be deducted from incomes of purchasers (if they are individuals and families); or an additional amount may be credited to payments by enterprises to individuals.

Each item in family income that cannot be included in national income (because it is not a return on productive activities) could be treated similarly. But such adjustment may seriously distort the concept of individual or family income, rendering it completely unusable for the study of the effects the amount of income received has on its disposition, on activities, or on non-overt states of families or similar units.

¹¹ Of course, a size distribution of individuals' income can help to explain how the size distribution of family income arises in the same way as a size distribution of national income can help to explain how differences in income (income more comprehensively determined) arise among various units. But whereas the types of income receipts that are properly components of national income constitute by far the preponderant part of income comprehensively determined for the purpose of studying size distributions, the difference between the individual and the family unit is far greater and the transition from the former to the latter much more difficult. Yet distributions of individuals' income are needed when analytical considerations of income earning are important.

Our failure to discuss here size distributions of individuals' income implies the absence of problems in defining the unit, not that such distributions are

useless. An obvious reason for not paying much attention to the individual unit is the absence of range in its definition and hence of difficulties in formulating it.

¹² This discussion raises an interesting question concerning the meaning of 'gifts' included in family income in many studies. The inclusion under income of gifts from individuals who, because of blood ties, may be expected to contribute, is partly due to the narrowness with which the family unit is defined.

¹³ In a frequency distribution of families by size of income each family is given the same weight. Unequal weighting of families is, of course, possible but gives a frequency distribution of other units of count, not of families.

¹⁴ This example suggests the danger of misinterpreting size distributions of family income when receipts are cumulated for periods as short as a year. For many purposes income derives its significance from its flow for much longer periods, and when we do not know what that has been, it is too easy to assume that income for a given year represents what might be called the family's income 'status'. But a distribution of *annual* income exaggerates differences in income status, as is evident from the presence of large groups of families with negative and zero incomes or incomes so small that by no stretch of the imagination could they be considered possible approximations to income levels for a substantial period.

Naturally, the inclusion in annual income distributions of transient sources of variability exaggerates also the number of families with exceptionally high incomes. When the concern is, as it often is, solely with family incomes below a so-called living standard, the effect of the unconscious identification of an annual income with income for a longer period is an exaggeration of the number of families whose income status is below a desirable minimum and the inclusion among the 'poor' of many families far from the poverty line (unless we specifically segregate groups whose annual incomes are likely to represent income status for a longer period). There is a similar bias when the number of the very 'rich' is emphasized, the number again being based upon income for a single year.

Of course, within certain limits such bias is valid. If a family receives in a given year an income of one million dollars, its income status is without question far above the average level for the country. And if a wage-earner family receives in a year of prosperity and under conditions of full employment (for the family) an income of only \$500, its income status is doubtless below an acceptable minimum.