By: L. E. Rowebottom and D. H. Steinthorson, Dominion Bureau of Statistics

The process of measurement has been defined as "an expression of the difference in distinguishable qualities or characteristics".1/ The first portion of this definition -- "expression of the difference" -- points to the "how" of the measurement process; the latter -- "distinguishable qualities or characteristics" -- draws attention to the "what" of measurement. Thus, defining what is to be measured is an essential part of measurement. This principle is well known and has important implications for the role of the statistician -- implications which have received considerable recognition in statistical literature and practice. The thesis of our paper is the proposition that in the field of economic and social statistics these implications deserve further consideration and still more emphatic recognition in the planning of surveys and the publication of survey results.

In fact, we would go so far as to claim that a significant gain in comprehension would follow from introducing a concept called "definitional error", analogous to sampling error and observational error. If it is reasonable to take "statistical error" in its broadest sense to mean the degree to which statistics fail to serve their scientific purposes, then it may be useful to consider choice of definition as a significant contributor to statistical error, worthy of theoretically equal status with other types of error.

The explanation of what we mean by definitional error is inherent in the proposition that definitions ought to be suited to the uses of statistics. It follows from this, that to the extent that the definitions employed in a particular statistic are not appropriate to an intended use, their use for the envisaged purpose involves error. Thus, if a particular series of economic or social statistics fell into disuse, both sampling and non-sampling error would continue to exist and statisticians could still debate, perhaps profitably, the nature and extent of these errors of measurement. But in such circumstances definitional error would be nonexistent since there would be no possibility of mistaken application. It is through use that definitional error comes into existence. Since our concept of definitional error is so closely allied with use, it could be argued that a more appropriate title would be error of application. However, we rejected this title because it seems to place the onus for controlling it on the user, whereas in our view statisticians have a definite responsibility for minimizing this type of error.

We realize, of course, that operational definition, although limited by what it is possible to measure, should be chosen in such a way as to ensure that what is in fact measured conforms as precisely as possible to the definition of what is to be measured. In fact, definitional error straddles the whole process of economic and social measurement from determination of concept, to formulation of questions, through tabulation and publication. However, we have considered the problems of definition involved in the enumerative process as a part of observational error outside of the scope of this paper. We will later comment on what we consider to be some of the definitional aspects of tabulation and publication. Thus, what primarily concerns us here is the relationship between the definition (or concept if you prefer) of what is measured for a particular purpose, and related definitions (or concepts) that might be used for the same or different purposes.

It is fruitful to examine the various processes whereby particular definitions come into use. In textbook parlance, the problem is simple and the procedure clear cut. The user provides a precise definition of what is required, including an attachment of permissible sampling error, and the statistician measures accordingly. It may be likened to the butcher filling an order of three pounds of sirloin steak, with, if you like, the requirement of "well trimmed" left to the statistician's judgment, perhaps subject to the customer's approval. In our experience this story book description varies from the real world in a variety of ways.

I. The user knows what he wants in a vague rather than precise way. He will likely know his purpose with precision, but not what statistics are required to achieve his objective. Occasionally this uncertainty provokes criticism of users whom it is felt should know exactly what they want. Such complaints are based on a misunderstanding of the field of knowledge of most users of economic statistics, who are primarily concerned with hypotheses to be tested or decisions to be made, and on a failure to appreciate the role that the statistician should play in assisting the user to decide what he wants. Users are concerned with such questions as controlling inflation, increasing productivity or sales, reducing unemployment or costs. When in order to make decisions they require measurements of prices, production or employment, it is unrealistic to expect them to know, as a matter of course, exactly what measurements they want, or to appreciate the vital connection between the definition of the requirement and its measurement. Because the statistician is involved in measurement and thoroughly conversant with its problems, he should be in a strategic position to assist the user in clarifying the nature of the measurement required. For example, consider what the

182

relationship should be between the businessman or the administrator who wants to know for some specific purpose, how many fishermen there are in British Columbia, and the statistician familiar with problems of defining occupation. To the user, "How many fishermen?" is likely to be a simple question requiring a simple answer. To the statistician accustomed to the difficulties of abstracting from reality the question is not simple, and before an answer which is not misleading can be provided, he must have a definition which will at least enable him to classify a person engaged in both fishing and farming. At the risk of belabouring the obvious still further, but we suggest not "obvious" to most recent graduates in statistics, it would be a worthwhile exercise to contrast (at greater length than we can do here) the relationship between statisticians and their clients, and that between architects or doctors and their clients. The latter professions do not make assumptions that clients have the knowledge to make decisions. However, we suspect that economic statisticians not infrequently do, without probing to see whether or not the assumptions are justified. We cannot resist making the facetious but perhaps illuminating comment that long ago society recognized that self-medication could be dangerous to the health of individuals, but not so in the case of statistics.

2. Textbooks generally describe the situation where the user is in a position to create the particular statistics he wants. In contrast, most economic statistics are the products of government and it is seldom that users are in a position to have produced the particular statistics they require. This is not because government agencies are unreceptive to the problems of users, but rather because some related statistic already exists and the modification required is impossible, too expensive, or unique to the particular user and not in the general interest to the point where additional expenditure is justified. Existing statistics must be adapted to the problem in hand, and the role of the statistician in evaluating differences between existing definitions and those required may be even more important than in establishing original definitions. It seems fair to generalize that the adaptation of existing statistics offers greater scope for definitional error than does the development of new statistics.

An example of this type of error arises in the use of city rent measurements. We are frequently asked to provide average rents for a particular city, and discover on questioning that the enquirer is concerned with what rent he may have to pay on moving. Quite aside from the question of applying averages to the problem of an individual, which could be partially met by the provision of frequency distributions, the enquirer has understandably failed to distinguish between the definition of rent relevant to his problem, and the definition of rent on which the statistics are based. In the latter case, the definition is rent paid by tenants for occupied accommodation; our friend requires a definition yielding a measurement of rent asked for vacant accommodation. One only needs to reflect on the difference between rent paid and asked in New York City where rent control is still in effect, to appreciate the possible extent of the dis-service involved in unquestioning provision of the statistic requested. In this case, the service of the statistician is to dissuade the enquirer from using the statistics and refer him to data which conform to the required definition -- the classified sections of daily newspapers.

3. Textbooks seldom describe the situation in which the statistician is not in close touch with many users who select statistics from published documents, and is compelled to make decisions on their behalf. While in such circumstances the statistician will, perforce, and with whatever consultation is possible, select his definitions in accordance with what he considers to be the predominant use, he may be almost certain that the statistic will be used to serve a variety of purposes impossible to anticipate. In these circumstances, it is particularly incumbent upon the statistician to provide precise statements of definition, and whenever possible statistics according to several definitions.

4. In final contrast to the textbook is the happier circumstance in which a set of relevant statistics applicable to a continuing problem has been in existence for some time. Here the subject matter expert and the statistician are thoroughly familiar with each other's problems, and the continuing improvement of definition and measurement may be described as a joint undertaking. However, even in this situation it has been our experience that there is room for improvement. As it was put by one of our colleagues who shall remain nameless, "I have seldom participated in an entirely satisfactory discussion with users". We wonder how many users would substitute the words, "with statisticians"?

It is from such diverse relationships between producers and users of economic statistics that definitional error arises. Even in such a variety of circumstances covering many subject matter areas it seems possible to offer some generalizations about the problem of definition in the field of economic and social statistics. First, economic phenomena are such that a variety of definitions is usually possible. Prices, productivity, production, and employment are complex subjects which defy representation in terms of single definitions, and frequently one would prefer not a particular defininition, but a comparison of results obtained from a variety of definitions. Second, within some range of definitions the user will almost always be indifferent, on the assumption that resultant differences in the statistics will be so small as to not affect decisions. Beyond this range he assumes differences to be significant, and will insist on the requirement of a particular definition.

4

Whether or not in a given case the classification is correct as between important and unimportant definitional differences, must await the statistical evidence. Thus in the field of price statistics a user may be indifferent as to whether or not the population covered by a consumer price index has an upper income cut-off of \$6,500 or \$7,000, but have strong convictions about the exclusion of farmers. Third, even in those instances where a satisfactorily unique definition is possible, the statistics serve a variety of purposes. While this difficulty may not exist for private agencies producing statistics for individual clients, it is most certainly the case with governmental agencies producing what have not incorrectly become known as multi-purpose statistics. In this case the words themselves suggest a variance with accepted statistical practice which states that the statistics shall be designed to measure a particular phenomenon and intended for a particular use. This cardinal principle is, of course, based on the obvious premise that different definitions which will produce different results, and not, as implied in the words multipurpose, that a particular statistic will satisfactorily suit a variety of purposes.

These aspects of definition -- the preference for more than one, indifference to some and concern with others, the multiple use of single statistics -- have implications regarding measurement of definitional error. By measurement of definitional error we mean, of course, determination of the differences between the results of statistical surveys which differ in definition but are otherwise the same in all respects. While it is seldom impossible to know beforehand the magnitudes of definitional differences, and consequently definitional error, nevertheless, to the extent that survey and calculation techniques incorporate a number of definitions, measurements of the magnitudes becomes possible. When such measurement is made, improved awareness follows on the part of both the statistician and the user. Statisticians become more aware of the significance of the definition employed, and users of the potential or realized error attributable to definition.

We believe that measurement of definitional error should be facilitated by continuing effort on the part of statisticians to incorporate a variety of definitions in surveys. We are convinced that understanding of economic events is enhanced by measurement according to more than one definition, and that publication of such results may commonly have a more salutory influence on uses, than publication of sampling error and sources of non-sampling error. This is because measurements of definitional differences provide users with alternatives from which it is possible to select the one most appropriate to the problem in hand, whereas statements of sampling and non-sampling error inform the user of particular aspects of the survey results which he cannot escape if he is to use the statistics at all. Of course, because of limitations of funds,

techniques or data, it is impossible to incorporate all useful definitions in any particular statistic and choice between alternatives is both inevitable and necessary. As among alternatives, predominant use and practical survey possibilities will determine the choice. However, it is our contention that, as a general rule, the greater the number of definitions which can be incorporated the more useful will be the results, and the greater the discrimination in use.

Take as an example statistics of farm income and size by type of farm. In attempting to define a farm as being of a particular type, say wheat, it becomes immediately apparent that there is no concensus as to what is a wheat farm. Some say that if fifty per cent of the total sales of a farm is composed of wheat, the farm should be so defined, others say seventy per cent, others a still different percentage. In such circumstances the most useful procedure is surely to type farms according to a number of definitions and let the user fit the statistics to his concept and purpose. It is worth noting the statistics themselves are unlikely to have much impact on the improvement of concept and purpose if the user is confined to one definition.

In the above example, farm income and size are likewise subject to a variety of definitions and in a recent Canadian farm income and expenditure survey, income is being defined in at least ten ways. It is our intention to publish results according to a variety of definitions of farm, income, size, and type. Thus we do not propose to publish an estimate of "the" number of wheat farms in Canada, but rather estimates that there are between x thousand and y thousand wheat farms, depending upon definition. Such statistics will force users to recognize the existence of the problem of definitional error, and as between alternative definitions will enable them to select the statistics most appropriate to each purpose.

It is sometimes argued that the publication of statistics in such variety is confusing to users and that the statistics should be left simple. As will be obvious this is a point of view with which we do not sympathize. Economic events are complicated and while statistics cannot reflect all of the complexities, it is our view that the user should not be misled, by over-simplification of statistical presentation, into believing that he is getting a simple answer to a simple question, when in fact he is getting a simple answer to a complex ouestion, and should be getting a variety of answers which illuminate the complexity.

Perhaps understanding the definitional error by both users and statistical practitioners would be furthered by greater attention to this subject in academic courses in statistical methods, which (as is occasionally pointed out) are too frequently designed exclusively or primarily for the research worker who will be doing his own statistical work. How can the implications of definitional error be taught? Since generalization 184

alone cannot carry us very far in the treatment of definitional error, its implications need to be taught by the case method, that is, by selecting a variety of subject matter fields to demonstrate the relations between uses and choice of definition. We applaud, therefore, the remarks of the three speakers at this Association's session last year on "Desiderata for the Basic Course in Economics and Business Statistics", in particular, the proposed undergraduate course outlined by George F. Break.2/ This course opens with a discussion of the concept of personal incomes and its uses, and goes on to relate statistical methods to uses of these statistics and the various available sources of data. Such training, we feel, would be helpful

in preventing the occurrence of some of the harm which can come from unrecognized definitional error.

REFERENCES

- I. Richard Stone, <u>The Role of Measurement in</u> Economics.
- 2. American Statistical Association, Proceedings of the Business and Economic Statistics Section, (Session of December 1958), 329.