UNITED NATIONS

ECONOMIC AND SOCIAL COUNCIL

GENERAL

E/CN.3/Sub.1/30
11 September 1950

ORIGINAL: ENGLISH

STATISTICAL COMMISSION

SUB-COMMISSION ON STATISTICAL SAMPLING

Fourth session

DRAFT REPORT TO THE STATISTICAL COMMISSION ON THE FOURTH SESSION OF THE SUB-COMMISSION ON STATISTICAL SAMPLING, 5 SEPTEMBER TO SEPTEMBER 1950

TABLE OF CONTENTS

INTRODUCTION

1. The Sub-Commission on Statistical Sampling held its fourth session from 5 September to September 1950. The Sub-Commission re-elected Mr. P. C. Mahalanobis as Chairman.

2. The following members were present:

Mr. P. C. Mahalanobis (Chairman)
Mr. W. E. Deming
Mr. F. Yates

The specialized agencies were represented by:

International Labour Organisation:	Mr. A. A. Evans
Food and Agriculture Organization:	Mr. P. C. Tang
World Health Organization:	Mr. G. E. Hill

3. The following representatives of Non-Governmental Organizations attended:

Category A:

International Chamber of Commerce:	Mr. G. L. Ridgeway

4. Mr. W. R. Leonard, Director of the Statistical Office of the United Nations, acted as the representative of the Assistant Secretary-General for Economic Affairs. Mr. W. J. Bruce, Assistant Director of the Statistical Office of the United Nations, served as Secretary to the Sub-Commission.

5. The following statisticians attended the meetings of the fourth session of the Sub-Commission and took part in the discussions:

Mr. E. P. Billeter	Statistisches amt der stadt, Zurich
Mr. R. H. Blythe	U.S. Air Force
Mr. R. Goodman	Survey Research Center, University of Michigan
Mr. M. Hansen	U.S. Bureau of the Census
Mr. W. G. Madow	University of Illinois
Mr. C. M. Mottley	U.S. Department of the Air Force
Mr. Y. Morita	Statistical Bureau, Japan
Mr. W. A. Shewhart	Bell Telephone Laboratories
Mr. R. T. Smith	U.S. Interstate Commerce Commission
Mr. R. Wilkinson	Bell Telephone Laboratories
Mr. S. S. Wilks	Princeton University
Mr. J. Van Yzeren	Central Statistical Office of Netherlands

6. The Sub-Commission adopted the following agenda for its fourth session:

A. Constitution and membership of the Sub-Commission.
B. Comments on syllabuses drafted in the third session.
C. Operational research
D. Sampling in tabulation.
E. Sampling to measure goods traffic by road,
F. Sample surveys of current interest.

G. Sampling methods for estimating distribution by size of individual and family income.
H. Methods to collect migration statistics.
I. Crop statistics.
J. Other subjects.
   a. Theory on extreme values and its technical application.
K. Report of the fourth session to the Statistical Commission.

CHAPTER I

CONSTITUTION AND FUTURE ACTIVITIES OF THE SUB-COMMISSION

1. The Sub-Commission had before it the request of the Statistical Commission to discuss the constitution and membership of the Sub-Commission.

2. In its report of the Fifth Session, the Statistical Commission observed that "the Sub-Commission has hitherto been most successful when the sampling problems on which its advice was sought have been submitted in detail and with full documentation"; and continued: "Arrangements are needed by which specialists in the subject fields in which the Sub-Commission is considering the use of sampling methods may participate more directly in the work".

3. The Sub-Commission fully agrees that sound advice on the application of sampling methods in specific fields can only be given when full information on the material to be sampled and the problems involved is provided, and it has always insisted on such information. It has also, wherever possible, arranged for the attendance of specialists in the particular subject field under discussion. It also recognizes, and has frequently stated, that the details of the most appropriate sampling scheme to deal with any particular situation can usually only be worked out by a sampling expert who is in contact with the preliminary stages of the work, and who is in a position to carry out investigations into the accuracy of the various possible methods, and if necessary can institute a pilot survey.

4. From the outset it has been recognized that the Sub-Commission from its constitution and because of the other committments of its members, has not been in a position to undertake this last service, which has, therefore, been regarded as the responsibility of the statisticians in charge of the survey. Indeed, the Sub-Commission has continually stressed the need for the training of statisticians in the methods of sampling.

5. In addition to its work on the application of sampling methods in specific fields, the Sub-Commission has concerned itself with the more general aspects of sampling methods, and as instructed in its terms of reference, has endeavoured to "promote the improvement of methodology in statistical sampling and its applications". To this end a standard terminology was drawn up which has not only served its direct purpose, but has also resulted in considerable

clarification of the logic of the subject. Two of its members have produced text-books on sampling, one of which would certainly not have been written had it not been that the need for such a book was expressed by the Sub-Commission. In the opinion of the Sub-Commission this work has been just as successful as the work in more specialized fields referred to by the Commission, and indeed has laid an essential foundation for work in specialized fields.

6. In the opinion of the Sub-Commission, therefore, it would be a grave error to attempt to confine the work to specialized applications of sampling methods, to the exclusion of such general problems as appear to it to be of relevance. The Sub-Commission has, in fact, been considering, in conjunction with the discussion on operational research reported below, whether the scope of its work should not be widened somewhat so as to include as its subject matter other statistical techniques which in its opinion are likely to be of value to members or organizations of the United Nations and its specialized agencies. It has come to the conclusion that this would be advantageous for the following reasons:

(a) Some of the existing and developing statistical techniques other than sampling are likely to be of value to members or organizations of the United Nations and its specialized agencies.

(b) The United Nations and its specialized agencies have at present no body which can advise it on the application of statistical techniques, other than sampling.

(c) The constitution of the Sub-Commission is at present such that it is well qualified to deal with such problems, as it is composed of statistical experts whose collective experience covers a far wider field than that of sampling methods.

7. The Sub-Commission, therefore, wishes to recommend that it be authorized to undertake work on these lines. It is not anticipated that at the outset the amount of such work would be very large. If the volume increases to such an extent that sub-division of the work of the Sub-Commission appears desirable, this can be considered when the need arises.

8. If this recommendation is accepted it suggests the following revised terms of reference be adopted.

/Terms of

Terms of Reference

9. The Sub-Commission on Statistical Techniques shall assist the Statistical Commission to promote the development and use of such statistical techniques as are likely to be of value to the Members of Organs of the United Nations, the Specialized Agencies and Non-Governmental Organizations brought into consultative status with the United Nations.

In particular, it shall:

(a) advise on the use of sampling methods in the collection of information needed by Members of Organs of the United Nations, etc., and make recommendations on the methods they consider most appropriate to specific fields under the varying circumstances existing in different countries;

(b) keep under review the use that is being made of statistical sampling in different countries, and in different fields of subject matter;

(c) determine in what circumstances sampling methods are preferable to complete enumeration in respect to reliability, speed, cost, continuity and other factors and make recommendations regarding the use of sampling methods in conjunction with attempted complete enumeration in order to assess the reliability of that enumeration, or to obtain early returns and supplementary information;

(d) bring to the notice of members, developments in other statistical techniques which appear likely to the Sub-Commission to be of use in the solution of problems which are directly the concern of Members of Organs of the United Nations, etc., and make recommendations regarding their use in specific fields;

(e) promote the improvement of the methodology of statistical sampling and other statistical techniques which are the concern of the Sub-Commission, and their applications.

Membership

11. The Sub-Commission's membership was originally set at 9 members but there have in fact never been more than 5 members (including 1 consultant). Moreover, since members attend in their personal capacity no substitutes are available in case of absence. It is considered that the membership might with advantage be expanded somewhat. It is recommended, therefore, that additional experts

/be asked

be asked to serve as consultants for one year in the first instance.

12. To make provision for the participation of subject-experts, as recommended by the Statistical Commission, it is recommended that the Sub-Commission be authorized to invite such experts to attend meetings in the capacity of temporary consultants in cases in which this appears advisable.

/CHAPTER II

CHAPTER II

COMMENTS ON SYLLABUS AS DRAFTED IN THE THIRD SESSION

1. Because of shortage of adequately trained statisticians in sampling, the Sub-Commission in its reports had stressed the need for training statisticians in sampling. Requests had been received from specialized agencies in regard to the development of a training programme. It was felt that the Sub-Commission could broadly indicate the level of training which should be expected of a statistician qualified in the field of sampling. With this object in view the Sub-Commission wished to stimulate discussion on suitable syllabuses for courses in sampling. At its third session two members of the Sub-Commission prepared syllabuses and the Statistical Office was requested to circulate them to appropriate organizations and persons for their comment. It was realized that while the primary responsibility for drawing up definite syllabuses or conducting courses on suggested lines must lie with universities, institutes and government and private organizations, the Sub-Commission might encourage further consideration and activity in this direction, especially when no generally recognized suggestions regarding syllabuses had hitherto been made. It may also be noted that the Sub-Commission did not endorse any of the syllabuses but only requested their circulation to obtain views from persons interested in education and training in sampling.

2. The large number of comments received by the Statistical Office clearly indicates that the suggested syllabuses had been successful in creating widespread interest. Thirty-one statements have been received from statisticians or groups of statisticians from 17 countries and most of these are useful and interesting. As the Sub-Commission does not propose to draw up sets of specific syllabuses, it will not pursue the matter further at this time.

3. The Sub-Commission, therefore, requests the Secretary-General to transmit the suggested syllabuses and the comments to the International Statistical Institute and to express the hope that the Institute will give further attention to this subject. The Sub-Commission would be pleased to receive from the Institute reports of further progress on this subject.

/CHAPTER III

CHAPTER III

OPERATIONAL RESEARCH

CHAPTER IV

SAMPLING IN TABULATION

/CHAPTER V

CHAPTER V

THE MEASUREMENT OF GOODS TRAFFIC CARRIED BY ROAD

1. Most countries collect and publish statistics on the movement of goods by rail, by inland waterway, by sea and by air; the statistics range from total figures like total tonnage loaded in the country to detailed statistics which analyse the total figures by commodity, by origin and destination and so forth. But the carriage of goods by road is performed by so many different types of carrier, including many small ones, that it has in most countries so far not been systematically measured despite the fact that in many placed road traffic is seriously competing with the railways.

2. The Economic Commission for Europe of the United Nations (ECE) has drawn attention to the importance of securing data on road traffic. The Inland Transport Committee of the ECE at its fourth session

> Noting (a) the increased importance of road transport in Europe, (b) the almost complete lack of statistics relating to this form of transport, particularly in respect of the volume of freight traffic and (c) the fact that the absence of road statistics renders it increasingly difficult to obtain any knowledge either of the total volume of transport as a whole or of the manner or direction in which transport is developing;
>
> Invited Governments..... to forward their views as to what may be considered the best methods of collecting the information necessary for the compilation of road traffic estimates, with particular reference to spot checks and samplings, giving as much information as possible on past experience in this field. (Excerpts from E/ECE/TRANS/166, Resolution No. 64)

3. Therefore the Transport Division of the ECE has co-operated with the Statistical Office in bringing the question of the use of sampling to measure goods traffic by road before the Sub-Commission on Statistical Sampling and in preparing documentation for the Sub-Commission (E/CN.3/Sub.1/24). The Sub-Commission also had available to it the documents listed in paragraph 1 of ANNEX I to this report and ECE document W/TRANS/WP6/12 and its annexes.

4. At the suggestion of the Working Party on Transport Statistics of the Inland Transport Committee of the ECE the problem thus put before the Sub-Commission was, in order to reduce it to a more manageable size by treating only the most important aspect, restricted to the study of obtaining the following total

figures for goods traffic by road motor vehicles:

(a) vehicle-kilometres, loaded and empty separately

(b) capacity in ton-kilometres on the basis of kilometres actually run

(c) tonnage of goods loaded

(d) freight ton-kilometres performed.

5. On considering the matter the Sub-Commission decided

(a) to treat in detail the restricted problem of obtaining the figures listed above in countries where motor vehicle traffic is dense, and to make a formal recommendation upon it;

(b) to make less detailed remarks on other aspects of the problem of measuring road traffic.

A detailed study of the restricted problem is appended to this report as ANNEX I; the Sub-Commission's recommendation appears in paragraph 6 below; the less detailed remarks appear in paragraphs 8 and 9 below.

6. The Sub-Commission on Statistical Sampling has noted the desire of the governments of many countries to obtain figures for road motor vehicle transport comparable to those which are available for railways, inland waterways and air transport. The Sub-Commission therefore requests that: the Secretary-General to recommend to those governments that they undertake in 1951 experimental sample surveys to obtain, during some conveniently chosen period such as a month, for road motor vehicle traffic the following total figures:

(a) vehicle-kilometres, loaded and empty separately

(b) capacity in ton-kilometres on the basis of kilometres actually run

(c) tonnage of goods loaded

(d) freight ton-kilometres performed

the Secretary-General to draw to the attention of the governments the technical suggestions for making such surveys which are contained in ANNEX I to this report and to ask governments to communicate to him their experience in conducting such experimental surveys for consideration by the Sub-Commission,

the Secretary-General to provide, within the limits of his resources, advice on the details of sampling technique when asked to do so by a government conducting an experimental survey of the sort recommended by the Sub-Commission.

7. ANNEX I to this report was prepared by the Sub-Commission on the basis of an exploratory paper submitted by the Statistical Office. The annex provides in

some detail an outline of how the sample survey mentioned in the recommendation above might be conducted, discusses the most economical way to select a sample, lists the information required to enable the quantities referred to in the recommendation to be computed and describes how their final computation can be made. In the Annex great emphasis is placed on the importance of enlisting the co-operation of carriers and carriers' associations. The representative of the International Chamber of Commerce present at the Sub-Commission meeting expressed the encouraging view that the Chamber might through its Transport Commission be able to assist in assuring the co-operation of carriers' associations in countries undertaking surveys.

8. The Sub-Commission considered the broader aspects of road transport statistics going beyond the four fundamental figures they studied in detail. It noted that for simplicity the following matters had been omitted from detailed consideration:

(a) the measurement of passenger traffic by road;

(b) the measurement of the movement of goods in road vehicles drawn by animals;

(c) the determination, as is required for road planning, of the number and type of vehicles using certain stretches of road.

The Sub-Commission also noted the importance of analysing by commodity and by origin and destination total figures on the movement of goods in order to be able thoroughly to study the relations between road and rail transport. The representative of the International Labour Office emphasized the importance in this connexion of determining the cost of carriage of goods by road, including the labour cost.

9. The Sub-Commission considered the uses of road traffic counts, with and without stopping of vehicles to interogate drivers. It is of the opinion that in countries where road network is not dense, stopping of vehicles and interogating drivers at a number of highway points might be the easiest way to obtain the total figures discussed in paragraphs 4-7 above and in ANNEX I. The Sub-Commission emphasized the importance of traffic counts in planning improvements of the road network.

CHAPTER VI

SAMPLE SURVEYS OF CURRENT INTEREST

/CHAPTER VII

CHAPTER VII

SAMPLING METHODS FOR ESTIMATING DISTRIBUTION BY SIZE OF INDIVIDUAL AND FAMILY INCOME

1. At its fourth session the Sub-Commission gave consideration to the use of sampling methods for estimating distribution by size of individual and family income (E/CN.3/Sub.1/SR.40 and E/CN.3/Sub.1/SR.41). It had before it a report on the usefulness of these statistics for the study of various economic and social problems, including a description of conceptual and sampling problems arising in their compilation (E/CN.3/Sub.1/27).

2. It is obvious in general that reliance on sampling methods in compiling income tax statistics has several advantages. In addition to considerably reducing the costs of compilation, it yields results more quickly and permits the individual returns to be treated with greater care and by more highly qualified personnel. The Sub-Commission, therefore, wishes to draw the attention of governments now publishing income tax statistics to the usefulness of sampling procedures in this particular field. Moreover, since income tax data are often not very reliable for certain purposes owing to under-reporting and are subject to other limitations (for example, differences in the time-periods to which the individual returns refer, the treatment by the fiscal authorities of certain items such as capital gains and losses, depreciation allowances, allowances for dependents and life insurance, etc.), statistics obtained by sampling, instead of by complete enumeration, may be sufficient for most purposes.

3. Because of the confidential nature of the income tax returns, and also for administrative reasons, fiscal authorities may not be prepared to make the basic material available for statistical purposes. Sometimes the original returns are accompanied by carbon copies containing the main figures. This is better than using separate sheets to be filled in by the assessees. However, either method gives more than is necessary if sampling methods are to be used. A practicable method may be to have the fiscal administration copy the data from the returns included in the sample, on special forms which thereupon are processed to the statistical authorities.

/4. Since the

4. Since the statistics are derived from the returns received by the tax authorities, probability-sampling will be the appropriate technique to be used. If basic information permits, stratification according to criteria believed to be correlated to income should be used. For example, in those countries where the fiscal administration is decentralised, the districts may be sampled separately. Detailed instructions will be necessary to avoid that in taking the sample its basic principles are unintentionally violated. For example, in the case of persons who were late in sending in their returns, or whose returns have not yet been approved by the revenue inspectors, substitution by other returns should not be permitted, but replacement by the return of the previous year may be possible under appropriate circumstances. Where the returns contain information, for example, on the number of dependent children, or other important characteristics, such information may also be included in the inquiries.

5. Statistics of income tax usually refer to the incomes of individuals, but may include joint returns of husband and wife. Other methods than those described above must be used for obtaining information on the distribution of family income and here, also, the Sub-Commission wishes to draw attention to the usefulness of sampling procedures.

6. Very often surveys of family incomes will be combined with inquiries on other subjects, such as expenditure on goods and services, ownership of assets, etc. The definitions of "family" and "income" to be used will depend on the purpose of the survey. In many surveys the "household" and not the "family" will be chosen as the basic unit. It is also possible to confine the spending unit to those family members who pool their incomes for all major expenses. If this is done, the family members who do not pool their incomes will be treated as separate units. Family membership may be determined as of the interview date, or each person who was part of the family during a portion of the year may be included. If the first method is used, only the income of persons being members of the family as of the interview date will be included. Persons who started earning an income during the year, newly-formed families, and other partial-year spending units may be included in the survey, or they may be excluded altogether.

7. In defining the income concept to be used, attention will have to be given to the treatment of such items as income in kind, farmers production for own consumption, net rental values of owner-occupied houses, contributions to social

insurance and pension funds, family allowances of members of the armed forces, gifts, capital gains, proceeds from change in ownership of assets, etc. Understatements of income in sample surveys may be expected to occur, owing to inaccurate reporting and a possible bias in the average incomes of families from whom information cannot be obtained owing to absence or refusal to co-operate. Wherever possible, estimates of total income derived from sample surveys should be compared with independent estimates of total personal income, obtained as part of the evaluation of national income, after duly allowing for differences in the underlying income concepts. Such comparisons may be carried out separately for the various income components, such as total wages and salaries, interest, rent, dividends, etc. The accuracy of the income data may, under appropriate circumstances, also be improved through other means, for example by examining both income and expenditure of the families concerned, and also the changes in total assets between the beginning and the end of the year. If circumstances permit, other methods might be used for determining the percentage understatement of income in sample surveys.

8. Multi-stage area sampling may be considered the appropriate device for obtaining data on the distribution by size of family incomes. Owing to non-sampling errors in surveys of this type, the reliability of the results cannot usually be improved very much by increasing the sample beyond a certain size. Information with respect to the small proportion of the population in the high income brackets usually weighs heavily in the means, aggregates and distribution of aggregates collected in the survey. To improve the precision of the results, oversampling of the high-income groups may be a useful device.

9. The use of sampling procedures for collecting information on the distribution by size of family income is greatly facilitated if, for the population as a whole, recent information is available on the frequency distribution of families by composition and size, occupation, number of children of different age groups, etc. Such information may be obtained in connexion with a census of population. The Sub-Commission, therefore, agrees fully with the following suggestions, made by the Population Commission at its fourth session:

"It is desirable that the number of 'family households' be tabulated by at least the following characteristics of households:

"(i) Total numbers of persons in the household (showing the numbers of households of 1 person, 2 persons, 3 persons, etc., up to 9, 10 and over).

"(ii) number of persons in the household who are related to the household head (showing households having heads with no relatives in the household, those with 1 person related to the head, those with 2 such persons, etc., up to 9, and those with 10 or more such persons.

"(iii) number of children, under 16 years of age, in the household who are related to the head (showing households with no children related to the head, those with 1 child, 2 children, etc. up to 9, and those with 10 or more). Wherever feasible the tabulation should be made separately for households headed by a married man and wife and for other households".

10. The usefulness of statistics of the distribution by size of family incomes for purposes of economic analysis would be greatly increased if they are classified by composition and size, occupation, and other family characteristics.

11. The attention of the Sub-Commission has been drawn to the need for investigation into family income and expenditure in the under-developed countries. Such studies are highly useful in that they throw light on actual conditions in those countries, and in that they may be of assistance in the study of methods aimed at raising the standard of living. The Sub-Commission suggests that the Secretariat, in advising those countries and interested Specialized Agencies, give attention to the appropriate methods that may be used in carrying out investigations into family income and expenditure in those countries.

CHAPTER VIII

METHODS TO COLLECT MIGRATION STATISTICS

CHAPTER IX

CROP STATISTICS

CHAPTER X

OTHER SUBJECTS

The Sub-Commission also considered a short statement on the statistical theory of extreme values and its technical applications prepared by Mr. E. J. Gumbel. The Sub-Commission expressed general interest in the subject but felt that it was in no position at present to evaluate the theory or to consider its application. It was believed, however, that further discussion should be deferred until a fuller exposition was possible. To this end the Sub-Commission requested the Secretary-General to investigate whether or not this subject could be fully prepared for consideration at its next session, and if so, to make such preparation with the assistance of an outside expert if necessary.

ANNEX I

SAMPLING TO MEASURE THE GOODS TRAFFIC CARRIED BY ROAD MOTOR VEHICLES

1. References.

(a) Economic Commission for Europe: Report of the Working Party on Transport Statistics on its third session (E/ECE/TRANS/224) referred to below as WP.

(b) Economic Commission for Europe: Report of the Meeting of Statisticians held at The Hague on 24 to 29 July 1950 (TRANS/ST/4) referred to below as S.

(c) United Nations Statistical Office: "International Standard Definitions for Transport Statistics" (Statistical Paper M 8) referred to below as M 8. Transport statistical terms will where possible be used in the sense of these definitions.

(d) United Nations Statistical Office: "The Preparation of Sampling Survey Reports" (Statistical Paper C 1, revised) referred to below as C 1 rev. Sampling theory terms will be used in the sense of this paper.

2. The object of this study.

As explained in paragraph 6 of the Report to which this study is annexed, it is desired to measure approximately the following quantities for goods traffic carried by road motor vehicles:

(a) vehicle-kilometres, loaded and empty separately (WP, paragraph 24);

(b) capacity in ton-kilometres on the basis of kilometres actually run (WP, 24);

(c) tons loaded (M 8, 23a; WP, 24)

(d) freight ton-kilometres (M 8, 23e; WP, 24).

This study outlines a method of measuring these quantities by sampling and will take the view that tons loaded in commercial traffic (M 8, 17) and freight ton-kilometres performed in commercial traffic are the fundamental figures required for economic purposes. It will suggest that vehicle-kilometres and capacity ton-kilometres can be obtained as a by-product of the process used to obtain figures for tons loaded and ton-kilometres.

3. The "frame" on which a sample is based.

The frame (C 1 rev., 2a) is the set of documents on which the choice and utilization of a sampling method is based. If possible the frame should be based

/on documentary

on documentary material already available in the country in question. This material may refer to:

(i) the carrier (permits to carry, permits to buy petrol, reports to supervisory authorities, insurance documents, etc.);

(ii) the vehicle (census, registration, insurance documents, etc.);

(iii) the goods (freight documents);

(iv) the locale where the carriage takes place (road maps, etc.).

The nature of the available documentation must determine the sampling procedure including the choice of elementary unit, sample unit, sample, etc. (C 1 rev., 2b, c, d).

4. The choice of a sampling method.

In this paragraph the choice of a sample based on each of the four types of material listed at 3 above will briefly be discussed. It is not intended that the discussion shall exhaust all possible uses of the material which may be available.

(a) Sample of vehicles. Because of national registration laws, a frame based on individual vehicles is probably easiest to construct and a stratified (C 1 rev., 2h) sample then easiest to set up. The operator of each individual vehicle, v, selected would be asked to report on the performance (loads handled, kilometres run etc.) of v in the time period P (say a month) covered by the survey. But in most cases the operator of v will not be able to report on the performance of v during P unless he knows at the beginning of P that he will be asked to do so. Such knowledge would, of course, introduce bias into the sample in a way which, while probably not serious in the case of operators of single vehicles, might be serious where the vehicle in question was one of a fleet operated as as a unit. For instance, if an operator owning three identical vehicles was told at the beginning of P that he would be required to report on a specific one of them, and if, for a part of P he had only work enough for two vehicles, it would embarrass him to know to which of his three vehicles to assign the work. The objection just described is here considered sufficiently grave to make it undesirable to base the survey on a straightforward sample of vehicles. It will, however, be recommended at (d) below that in certain cases a sample of carriers be based on a sample of vehicles.

/(b) Highway traffic

(b) Highway traffic counts. The frame for a highway traffic count can probably be constructed in any country but in countries where the road network is at all dense a traffic count is so cumbersome and expensive that it should not be undertaken unless the paucity of documentary information of other sorts makes a traffic count necessary. It is here assumed that this case will not arise in any European country. The situation in countries where traffic is not dense is briefly discussed in paragraph 9 of the Report to which this study is an annex.

(c) Sample of freight documents. Freight documents may be in the hands either of shippers, carriers or consignees. The multiplicity of shippers and consignees is so great that it is probably impossible to obtain lists of either of these from which samples can be drawn. So if freight documents are to be used they must be obtained from carriers. It is probable that operators of large fleets use freight documents from which can be obtained most of the information required about goods handled, and advantages will below be taken of this fact (8, e). But small operators, particularly those who operate single vehicles, are unlikely to make systematic use of frieght documents and it therefore appears to be undesirable to attempt to sample freight documents directly (compare 8, 28).

(d) Sample of carriers. A sample of carriers, including carriers for hire (M 8, 15a) and private carriers (M 8, 15b) seems to offer the fewest objections. Each carrier sampled would, as described in paragraph 8 below, be asked to report the total tonnage of goods he has loaded in a period, I, of about a month inside the country in question and to supply some additional data. It is recognized that carriers, particularly large carriers, are likely to resist providing information from which their total revenues could be estimated. But because of varying rates for different commodities the figure for total loadings would not, in general, make it possible to approximate revenue. Where serious objection is nonetheless feared, the suggestion made in paragraph 8 e below may be of some use. Paragraphs 5-11 below discuss in some detail the sampling of carriers.

5. Listing and sampling carriers.

(a) It is unlikely that in any country a complete list of carriers is available which includes all carriers for hire and private carriers no matter how small their operations. But partial lists may be accessible in

/the hands of

the hands of regulatory authorities, carriers' associations, insurance companies, etc. As many as possible of these should be assembled and the co-operation of the authorities, associations and companies should be enlisted (S, 29e). If a sufficiently complete coverage of carriers results from these lists, the carriers can forthwith be numbered serially and sampled by a random method. If information is available on the nature of each carrier's operations (S, 18) and the number of his vehicles, the sample can be stratified according to this information.

(b) If, as will usually be the case (S, 28), the lists discussed at (a) omit a significant number of carriers, the registers of the numbers (on license number plates) used to identify the lorries can be used to construct a sample of carriers (S, 27). The sampling procedure might be as follows:

(i) From the registers mentioned above the requisite number of lorries could be drawn, perhaps simply by picking out license numbers at random.

(ii) The sample of carriers would then be the set of those carriers who operate the vehicles selected at (i).

(iii) In computing the final figures allowance must be made arithmetically for the fact that the probability of selecting a carrier operating several vehicles is greater than the probability of selecting a carrier operating a single vehicle and that some carriers may be "selected" more than once.

(iv) In deciding on the size of the sample it should be remembered that a small sample, chosen truly at random, a large proportion of whose members co-operate in the survey produces more accurate results than a larger sample which is biased either in its initial selection or by a large amount of non-response from the carriers sampled. Methods of avoiding non-response are discussed at 8 f below.

6. Relative accuracy of the various methods

If the carriers can be stratified by nature and number of vehicles operated (5a), a sample of carriers, stratified according to this information with a variable sampling fraction (C 1, rev. 2i) approximately proportional to the size of operations, will provide the most accurate results. On the other hand, if the information on carriers is not adequate to effect this stratification a sample of

/carriers based

carriers based on vehicles (5 b) may be expected to give more accurate results than would a random sample of carriers, since the procedure of 5 b will sample carriers with probability approximately proportional to the number of vehicles operated. In cases in which the required information is available only for certain categories of carriers, a stratified sample of these categories can be taken, the remaining carriers being sampled by means of vehicle registrations, only carriers not occurring in the categories on which the stratified sample is based will be retained in the latter sample.

7. Analysis of the final figures by class of carrier.

The final figures can of course be analyzed, for instance by class of carrier (S, 18), even if they were not obtained from a stratified sample. To maintain a given accuracy in the figures for different categories the size of sample must increase with the number of categories; these must therefore be known before the sampling fraction is fixed. In particular it can be arranged that the survey shall yield to a desired degree of accuracy, the average performance per vehicle (overall or by class of carrier) thus relating the result of the survey to data on the number of vehicles in use, a type of data which is normally kept up to date in most countries (S, 29f).

8. The information to be collected from the carriers sampled.

(a) If P is the time period for which data are to be collected, the carriers to be sampled should be notified before the beginning of P of the questions they will, at the end of P, be asked to answer. This appears to be a necessary step if adequate replies are to be expected even though it may introduce bias into the sample. Since a report on total operations will be asked for, the bias introduced will be negligible compared to that referred to at 4 a above.

(b) The principal datum each carrier would be asked to give is the total gross weight (i.e. weight of goods plus their packing (M 8, 22b)) of the commercial goods loaded into his vehicles inside the country in question in period P.

(c) For a relatively short sub-period, Q, of P, say a day or two if P is a month, each carrier sampled should be asked to keep a "complete record" showing the following facts for each vehicle which he operates in the country during Q. The complete record should include:

(i) the license number of the vehicle;

(ii) the distance

(ii) the distance it has run in the country in period Q and the distance it has run empty;

(iii) for each consignment loaded within the country during Q, its gross weight, place of loading, the place at which it is either to be unloaded or to cross the frontier for unloading abroad;

(iv) For each consignment entering the country loaded (M 8, 23c) in the vehicle during Q, its gross weight, point of entry, the point at which it is to be unloaded or, if it is in transit, the point where it is to leave the country.

As operators may be expected to be reluctant to show overloads, some bias may be expected in the complete records so obtained. If, under (i) above, the rated capacity of the vehicle is asked for the bias will be increased so rated capacity should not be asked for if it can be obtained from other sources (see 5b, iv and v above).

(d) The periods Q should not run concurrently for all carriers sampled but should vary from carrier to carrier so that the whole of P is adequately covered.

(e) If some carriers in the sample regularly make use of freight documents, arrangement can be made through a carriers' association or by an interview (see f (i) and (ii) below) to permit them to submit a sample of these documents instead of the reports described at (a), (b) and (c) of this paragraph thus saving them clerical work and avoiding the necessity of any one of them divulging the extent of his operations. The carrier would have to number serially the documents he used in period P and submit the ones with specific serial numbers selected at random.

(f) Some suggestions follow for minimizing the number of carriers who do not co-operate with the survey.

(i) It should be determined from the lists mentioned at 5 a above which of the carriers in the sample belong to associations, and as much help as possible should be obtained from the associations in getting the information described above from those of its members who are in the sample. It would be helpful if the association itself would provide information on the number, license numbers and capacities of the lorries operated by the carriers.

/(ii) It is

(ii) It is desirable that carriers in the sample who do not belong to associations be interviewed with the object of enlisting their co-operation and obtaining the information mentioned in the last sentence of (i) above (S, 29a). The extent to which this can be done depends, of course, on the number of carriers in question and the funds available.

(iii) The number of carriers who do not co-operate can be substantially reduced by the methods described at (i) and (ii) but there are in any case likely to be some carriers who do not co-operate. Though the final figures can be adjusted to allow for them, there is no remedy for the decrease in accuracy for which they are responsible.

9. The data obtained from the survey.

The process described at 8 above will make possible the estimation of the quantities listed below.

(a) The total gross weight, L, of commercial goods loaded inside the country by domestic road motor vehicles.

(b) If "local traffic" is defined as the traffic hauled less than x kilometres, x being determined in each country to include in local traffic short distance traffic not competing with other means of inland transport, the complete records make possible the estimation of the ratio, f, of local to total traffic.

(c) From the complete records the average length of haul, h, of a ton of goods moving in domestic traffic can be estimated.

$h = \frac{\Sigma d_i x_i}{\Sigma x_i}$ where x_i is the weight of the ith consignment of goods and d_i is the distance it was carried.

(d) From the complete records the average length of haul, h', of a ton of goods moving in international traffic in domestic vehicles can be estimated.

(e) From the complete records an analysis could be made showing the quantity of goods hauled for each of a representative range of distances (S, 19).

(f) From the complete records vehicle-kilometres and the degree of loading can be estimated.

(g) From the complete records and the data on which the sample is based

/capacity

capacity ton-kilometres can be estimated.

10. International traffic.

If special arrangements can be made for a check of all vehicles at frontier points during the period P (see WP, 4-5), the following quantities can be accurately determined.

I_1 the gross weight of commercial goods imported in domestic vehicles

I_2 the gross weight of commercial goods imported in foreign vehicles

E_1 the gross weight of commercial goods exported in domestic vehicles

E_2 the gross weight of commercial goods exported in foreign vehicles

T_1 the gross weight of commercial goods moving in transit in domestic vehicles

T_2 the gross weight of commercial goods moving in transit in foreign vehicles

Where there is an appreciable amount of local traffic crossing the frontier it should, if possible, be excluded from the figures.

11. The computation of commercial traffic.

The quantities determined in paragraphs 9 and 10 can be combined as follows.

(1) Weight of goods loaded in long distance traffic:

$(1 - f) L - E_2$

(2) Weight of goods unloaded in long distance traffic:

$(1 - f) L - E_1 - I_1 - I_2$

(3) Weight of goods carried in long distance traffic:

$(1 - f) L - E_2 - I_1 - I_2 - T_1 - T_2$

(4) Freight net ton-kilometres performed in long distance traffic:

$h ((1 - f) L - E_1) - h' (E_1 - E_2 - I_1 - I_2 - T_1 - T_2)$

These formulas are subject to the following inherent errors in addition to the sampling errors.

(a) Duplications. Goods transloaded from one long distance lorry to another will, in formulas (1), (2) and (3), be counted each time they are transloaded. (Transloading from vehicles engaged in short-haul collection or distribution to long distance vehicles does not give rise to duplication because of the use of the factor 1 - f.)

(b) Omissions. Consignments carried in the domestic traffic of the country (i.e., both loaded and unloaded in the country) by vehicles of foreign registry will not appear in the figures.

Neither of the errors just listed is likely to be appreciable.
