

Project Title
Strengthening Statistical Capacity for the
ESCWA Countries in Energy Statistics
and Energy Balance

Meeting on Energy Balance and End Use in the Transport
Sector

22-23 December 2014

Department of Statistics, Amman, Jordan

Abdulhakeem Eideh
Al-Quds University, Palestine
22-23 December 2014

Methodologies for Data Collection of
Energy use of the Transport Sector

OUTLINE

- **Data Needs**
- **What should be collected?**
- **Who are the potential respondents - problems of co-operation?**
- **The Population**
- **Preparations for the survey Sample design**
- **Goods Road Vehicles**
- **Sampling Frame**
- **Sampling Designs**

Data Needs

Identifying user groups - who wants what

- **Government is often the main user group as well as the source of funds for the survey.**
- **Research institutes need the same data as government but may require more detail. These institutions often carry out work commissioned by others. For example, analysing the consequences of a policy measure or a proposed measure.**
- **Educational establishments' needs are very diverse but it is also very difficult to obtain a consensus view, as most users from this area tend to want very detailed data for a very narrow field of interest.**

- **Businesses want information about their industry group as well as data on general economic indicators.**
- **Trade associations and other non-profit organisations want similar information as businesses.**
- **The general public is an important market for statistics. It is difficult to gain a view of their demand for data, as their needs are widespread. One way to assess these needs is to maintain a note of all the requests for data in the subject field of a survey that had been received over the past three years.**
- **Although this manual concentrates on the data needs of the European Commission and the Member States, the needs of other international organisations should not be overlooked.**

What should be collected?

- **Allocate the variables to four groups:**
- **The essential data – if you cannot collect this data, then you need to go back to those who asked you to collect it in the first place for some very serious discussions.**
- **The desirable data - in this group comes data which is almost essential but you may need to give operators sufficient time to organise their recording systems to be able to provide the information to you.**
- **The useful data - helps interpretation but be careful not to overload the questionnaire.**

- **Asking too many “nice to know” questions may be a major reason for poor response rates.**
- **The "not this time" data - in this group will come the requests that do not readily fit into the general scope of the survey or are too expensive, too detailed, likely to cause offence to the respondents, etc.**
- **If extensive discussions have been held with a wide variety of user groups, it is likely that quite a number of the requests received will fall into this fourth group.**

Who are the Potential Respondents - Problems of Cooperation?

- **In most countries the road haulage industry is characterised by consisting of a large number of owner-operators - a single person with one lorry, or a small company with one or two lorries.**
- **At the other end of the scale, there will be one or two very large companies owning a large number of vehicles, possibly spread around a number of depots.**
- **Although around 80 per cent of the companies may be owner operators with only one or two vehicles, they are likely to account for only 20-30 per cent of the total road haulage activity.**

- **The 20 per cent of companies being larger will deal with 70-80 per cent of activity (measured in terms of tonne-kilometres).**

The Population

- **It is important to define exactly the area of interest – the statistical population.**
- **All goods vehicles (including small pick-up trucks and vans with a carrying capacity under one tonne)**
- **Or should there be a "cut-off", e.g. only goods vehicles with a carrying capacity of one tonne or more?**
- **The Regulation (EU) permits Member States to apply a "cut-off" of up to 3.5 tonnes carrying capacity or 6 tonnes gross vehicle weight.**
- **When carrying out surveys of goods road transport, countries should consider the structure of their road transport industry when deciding the exclusion level of small vehicles from their survey.**

- **For the first survey carried out by a country, it is often useful to include vehicles with a smaller payload than 3.5 tonnes (from 1 or 1.5 tonnes) in order to measure the amount of national transport performed by these small vehicles.**
- **Having measured the work of these small vehicles on the first survey, they can be excluded from future surveys if their work is only a small percentage of the total.**
- **Alternatively, the smaller vehicles might be included in a survey once every five years.**

- **Additionally, the Regulation does not apply to "vehicles whose weight or dimensions exceed the normal permitted limits of a country" or to "agricultural vehicles, military vehicles and vehicles used by public administrations and public services with the exception of road vehicles used by railway administrations".**
- **Information about the operator (i.e. the road haulage company), the activity (i.e. the carriage of goods) or the client (i.e. the freight forwarder or the customer)? Regional estimates as well as national figures?**

Preparations for the Survey Sample Design

Goods Road Vehicles

- **The journeys undertaken by goods road vehicles are sufficiently homogeneous that sampling can be carried out either of a vehicle journey or a vehicle time period (usually a week).**
- **It is important that sampling is carried out throughout the time period of interest - normally a calendar year.**
- **Almost certainly there will be considerable variation in the tonnes carried and tonne-kilometres performed between the very small rigid goods vehicles and the large articulated road tractors with semi-trailers.**
- **The sampling strategy should take this into account.**

- **In designing any sample, the following points have to be considered:**

1. Can you limit the burden on respondents:

- **By excluding some areas (for example, one-man businesses - not a good idea for road haulage)?**
- **By limiting the frequency any hauler is approached, or ensuring no vehicle is selected more frequently than once a year?**

2. Will stratification of the sample improve the efficiency?

Sampling in Space

- In order to use a sample, you have to ensure there are enough events occurring over the period you are covering of a sufficiently homogeneous nature in order to give you the sample estimates you want with sufficient accuracy.
- This may well depend upon
 - whether you want information about domestic
 - and/or international movements; whether you want regional and/or national estimates;
 - whether the respondents are owner-operators or companies,

- and, if companies, whether they are carrying out own account or hire or reward transport.
- The information about the carriage of goods by road is collected basically from the loaded journeys made by goods road vehicles
- For most countries, the vehicles making these journeys are sufficiently numerous for a sample of these vehicles to provide sufficiently precise estimates of the required measures of activity.
- If there are 100 000 relevant goods road vehicles on a **national vehicle register** and a sample of 10 000 such vehicles are selected over a survey period of one year, then the sampling density in space would be 10 per cent.

- **However, in some cases it is not possible or it is inefficient to use a national vehicle register as a sample frame for the survey**
- **For instance, if international journeys form a very small percentage of all journeys, a sample of vehicles from the national vehicle register might not produce a sufficient number of international journeys to provide the required activity measures with the necessary precision**
- **In such cases a register of operators authorised to carry out international transport might be used as the sample frame.**
- **Operators selected from this register would be asked to provide information about international journeys made by their vehicles.**

- **The number of operators on the register might be such that it would be necessary to approach every operator during a survey year. In such a case the sampling density in space would be 100 per cent.**

Sampling over Time

- **Even where the sampling density in space has to be 100 per cent,**
- **Which means that every sampling unit has to be approached during a survey,**
- **The journeys undertaken by goods road vehicles are sufficiently homogeneous that sampling can be carried out either of a vehicle journey or of the activity of a vehicle over a number of days.**
- **A time period can be chosen to limit the amount of information that would have to be returned for any one vehicle.**
- **The period normally chosen is for activity during one week.**

- **It does not matter which day is chosen to start a week, provided that the same start day is used throughout a year and each week comprises seven days.**
- **Where the choice is made for the recording of the activity of a selected vehicle over one week, the sampling density over time is 1.92 per cent (1/52).**
- **For international journeys the problem can arise that a vehicle may be on a journey that started before the week selected for that vehicle;**
- **The end of that journey may be during the selected week or even after that.**
- **Journeys may also start in the selected week and extend beyond the end of that week.**

- **The solution here is to provide instructions that the journeys to be recorded are all those that start in the selected week and to record the details of the full journey even if it ends after the selected week.**
- **Trade patterns and seasonal effects have to be taken into account when carrying out surveys.**
- **These are likely to be important in road haulage surveys.**
- **This means that to avoid the possibility of bias in the results, a survey of the carriage of goods by road should aim to cover activity throughout the survey period (that is, all 52 weeks of a year should be surveyed to produce estimates of activity for the year).**

Sampling over Domains

- **Almost certainly there will be considerable variation in the tonnes carried and tonne-kilometres performed between the very small rigid goods vehicles (lorries) and large articulated road tractors coupled to semi-trailers**
- **The sampling strategy should take this into account**
- **The best way to do this is to stratify the sample by the load capacity (or gross weight) of the vehicles**
- **Road tractors should form separate strata of the sample.**

- **Although they do not have a load capacity as such, in most countries the registration details for these vehicles include a maximum load capacity (or gross vehicle weight) that the tractor is permitted to tow**
- **If the weight data is not available for road tractors, they should still form a separate stratum of the sample.**
- **In larger countries it is also helpful to stratify the sample by region or by some other geographical division**

Avoiding Bias

- Bias in a sample can cause systematic, non-compensating errors that **cannot be reduced or eliminated by increasing the sample size.**
- Bias in sample selection can arise:
- If the sample is selected **other than randomly.**
- If the **sampling frame does not cover** the population adequately, completely or accurately.
- If some of the selected units **refuse to reply or cannot be contacted.**
- If some of the questions are answered incorrectly (**response errors**).

- **If the sample is selected other than randomly:**
- **Drawing the sample randomly is essential.**
- **Where information is available to allocate the units in the population into groups (strata) with similar characteristics,**
- **Drawing a random sample from each group does not destroy the randomness of a sample.**
- **In fact, the use of stratification is very beneficial if applied sensibly as it can lead to a significant improvement in the precision of the estimates.**

- **If the sampling frame does not cover the population adequately, completely or accurately:**
- **If the sample frame does not cover the population adequately, then the units which are not in the sampling frame have no chance of being included in the sample.**
- **For road vehicles, this is most likely to occur for newly registered vehicles.**
- **Hence, if a **vehicle register** is used as a sampling frame, the older the vehicle register is in relation to the date of the survey period of the sample, the greater the likelihood of bias.**
- **It is usual for the sample for a quarter of a year to be selected from a vehicle register about 6 weeks before the start of that quarter.**

- **The sample for the 13th week of that quarter will have been chosen from a register that is some 4 months out of date.**
- **If the sample had been chosen for a full year, the one for the last month of the year would have been based on a register that was over a year old - all newly registered vehicles during that year would have no chance of being included in the sample.**

- **Keeping response errors to a minimum depends upon good questionnaire design and good management practice in the survey office.**
- **Miscoding and data encoding errors are included as response errors although they do not arise through the respondents.**
- **The provision of false information by a respondent is very difficult to detect. Fortunately, for road transport surveys, it is probably more difficult for a respondent to create false information about journeys than to provide details of the actual journeys made.**
- **The greatest problem is that some respondents will claim that the vehicle has not worked during the survey week rather than record the actual activity, as this reduces markedly the work required.**

Registers for Road Transport surveys

- To draw a sample, we need a **sampling frame** –
- A register that lists all the units we wish to sample
- In this case goods road vehicles,
- And a contact address of the owner,
- Or a register that will enable us to sample these vehicles through **their operators**.
- For a **survey of the carriage of goods** by road, the best register will be that of the **goods road vehicles**.

- If this is not available or not sufficiently reliable, **a register of persons licensed to operate as road hauliers (company/registered owner for private hauliers) or a business register of companies could be considered.**

- **For each possible register the following questions need to be asked:**
- **Who produces the register?**
- **Can you gain access to it?**
- **Will the owners allow you to use it to draw a sample?**
- **Or will they draw the samples for you?**
- **How frequently will the owners allow a sample to be drawn - monthly, quarterly?**
- **(There may be legal restrictions against the use of some registers!)**
- **How good is the register in terms of quality?**
- **Is the information accurate?**

- **Is the information up-to-date?**
- **How long is it before new registrations are included - a month, a year, two years?**
- **Does it identify the vehicles you wish to sample or will your sample include vehicles outside your range of interest (e.g. cranes)?**
- **Cut offs? How comprehensive is the register?**
- **Does it exclude small companies?**
- **Does it exclude vehicles in certain categories (vehicles owned by public bodies, goods vehicles with a very low carrying capacity)?**
- **If there are deficiencies in the register you would like to use, can you work with the owners of the register to improve it?**

- **For a survey of the carriage of goods by road, the best register will be one that includes goods road vehicles.**
- **This is because the activity we wish to measure is the work that is done by goods road vehicles.**
- **That is, the weight and type of goods carried; the place of loading and unloading and the distance travelled;**
- **The characteristics of the vehicle and some information about the operator of the vehicle.**
- **Sampling vehicles also means that the burden of filling in questionnaires for operators with a large number of vehicles will be spread over time.**

- If a vehicle register is not available or not sufficiently reliable,
- There may be a **system of licensing road haulage operators and a register of persons licensed to operate as road hauliers** (company/registered owner for private hauliers) could be considered as a sampling frame.
- Using such a register means that the **sample of vehicles will be clustered** as the **operator** will be the **sampling unit** and thus the burden of questionnaire completion for operators with large numbers of vehicles will be concentrated into one time period - which could have an adverse effect on response

- Alternatively a **business register of companies** could be used.
- However, since this is a register of companies it has the defects of the licensed operators register.
- In addition it will exclude any operator whose business does not have to appear on the business register.
- In addition, the use of the business register as a first stage-sampling unit means that the main business of most of the companies selected will not be road haulage and many of them may not carry out even own account transport

- It is strongly recommended that, if at all possible, **a vehicle register should be used as the sampling frame for surveys of the carriage of goods by road.**
- Even if there are problems with the quality of data recorded on the register, a vehicle register is usually the only one available that lists directly the units (vehicles) which one wishes to sample.
- Throughout notes it has been assumed that the **vehicle register will be used as the sampling frame.**

- Nevertheless, despite the recommendation to use a vehicle register if at all possible, experience has shown that there may be **considerable problems** with their use as a sampling frame.
- **Firstly**, it is unusual for the organisation responsible for the maintenance of the vehicle register to be the same as that carrying out road haulage surveys.
- Road haulage surveys are normally carried out by the National Statistical Institute but the vehicle register is usually the responsibility of the police or another Ministry (often the Ministry of the Interior or the Ministry of Transport).
- Access to the register by the National Statistical Institute may require delicate negotiations at a very senior level in both organisations

- **Secondly**, despite any assurances to the contrary given by those maintaining the vehicle register, be prepared for errors in the data on the register and for it to be out of date.
- Vehicles may have been sold or scrapped within the last few months and the information not yet been added to the register.
- This may also have occurred between the time the sample was drawn and the questionnaire posted to the respondent.
- For vehicle registers that are used to collect vehicle taxes and have been running for many years, the number of such cases is likely to be very small.
- However, for registers that have only been set up for a few years, or are in the process of being created, experience has shown that the number of cases is quite considerable where the person contacted no longer possesses the vehicle.

- **Another problem with vehicle registers is that the address recorded may be insufficient for questionnaires to be delivered by post.**
- **If the vehicle register also includes a business reference number for the owner, it may be possible to amplify the address from the vehicle register by checking the address on the business register for that reference number.**
- **A further problem relates to the long-term hire of vehicles.**
- **The owner of a vehicle recorded on the vehicle register may hire it out to another operator on a long-term basis.**

- **By the time the survey team have received the information about the name and address of the true operator it is usually too late to send a questionnaire to the true operator**

- **What can be done to overcome these problems?**
- **To some extent a slightly larger initial sample of vehicles can be selected where it is expected that the register may be not completely up to date.**
- **However, care needs to be taken when grossing up the sample results to national estimates.**
- **The assumption made is that the raw survey results are representative of the total numbers recorded on the vehicle register.**

- If, from the sample drawn, say, 5 per cent of the returned questionnaires indicate that the vehicle has been scrapped, these returns need to be treated as valid responses since they indicate that around 5 per cent of the vehicles of that type on the vehicle register are actually scrapped.
- To exclude the scrapped vehicles found on the survey from the sample numbers when grossing up to national figures would result in an over estimate of the tonnage and tonne-kilometres performed.
- The calculation of weighting factors for grossing-up the survey results where registers may not be up to date is covered **in Counties Visits.**

- **Close liaison should be maintained with those responsible for maintaining the vehicle record.**
- **The experience where some of these problems have occurred is that, where the evidence of errors has been shown to those responsible for the vehicle registers, there has been cooperation and a willingness to take steps to improve the data on the vehicle register.**
- **However, it is likely to take some years to make a significant improvement to the quality of the data on registers where more than one in five of the entries prove to be in error**

Defining the Sampling Frame

- Having identified the best register to use to draw the sample for the survey and obtained the agreement of those responsible for the maintenance of this register that it may be used as a source for the sample of statistical units, **it is necessary to examine the data that is held on this register.**
- **The first step** is to make a list of the statistical units (vehicles in the case of a vehicle register) that should be excluded from the survey.
- What these exclusions are (for example, vehicles with a load capacity less than 1 tonne) should have been decided when the objectives of the survey were agreed.

- However, it may be that the register that will be used does not contain all the necessary information to permit the desired exclusion.
- This may lead to a review of the coverage of the sample
- **The second step** is to list all the information on the register that you would wish to capture to add to the survey record.
- If the register already contains all the data you wish to have about a vehicle, it is sensible to abstract that data from the register and include it on the questionnaire sent to the respondent.

- **The respondent then can be asked to check this data and correct it if it is wrong.**
- **This is much less of a burden than asking the respondent to write this information on the questionnaire.**

Sample Designs

- **Simple random sampling**
- **Stratified sampling**
- **Cluster sampling**
- **PPS sampling**
- **Two-Stage sampling**

S
T
O
p