

## ANNEX 3

# Sample Survey Design and Data Quality Control

### **An Example with Cameroon**

#### *Survey Coverage*

The Trucking Survey in Cameroon targeted trucking companies and companies conducting their own transportation. A trucking company is defined as a company that conducts trucking as its main operation and that has five or more full-time paid employees. The companies surveyed serve at least one of the following routes:

- Douala–Yaoundé
- Douala–Ngaoundéré
- Douala–N'Djaména
- Douala–Bangui
- Douala–Bertoua
- Douala–Garoua Boulai
- Douala–Bafoussam
- Yaoundé–Bafoussam

The survey also sampled a selection of truckers (trucking operators with fewer than five full-time permanent paid employees) that serve the main roads listed above.

### ***Companies with Five or More Full-Time Paid Permanent Employees***

A list of Cameroonian trucking operators was obtained from the World Bank Transport Unit. This list was completed and updated during the early stages of the survey. To validate the list and to classify the establishments according to size, the EEC team tried to contact a selection of establishments drawn randomly from the list. Following the results of the validation process, a sample frame consisting of a population of 52 establishments was set.

An attempt was made to contact each of these establishments. During the survey, it was discovered that 10 establishments were closed, 8 establishments were unreachable despite repeated attempts by phone, 12 establishments refused to participate, and 22 establishments agreed to participate, resulting in 22 completed trucking questionnaires, of which 5 companies provided their own transportation.

### ***Truckers***

In this survey, the trucker's stratum covers all establishments of the trucking industry with fewer than five employees. For many reasons, including the small size of establishments, their expected high rate of turnovers, the high level of "informality" of establishments, and consequently the difficulty in obtaining trustworthy information from official sources, the survey firm selected an aerial sampling approach to estimate the population of establishments and select the sample in this stratum according to the roads to be covered.

First, to randomly select individual truckers establishments for surveying, the following procedure was followed: (i) select districts and specific zones of each district where there are lorry parks or where truckers usually off-load; (ii) count all truckers who generally stop in these specific lorry parks; (iii) in accordance with this count, create a virtual list and select establishments at random from that virtual list; and (iv) on the basis of the ratio between the number selected in each specific zone and the total population in that zone, create and apply a skip rule for selecting establishments in that zone.

The survey firm went into the field to count truckers in the selected areas. Once the count for each zone was completed, the numbers were sent back to the firm headquarters.

At head office the following procedure was followed: the count by zone was converted into a list of sequential numbers for the entire survey region, and a computer program performed a random selection of the determined number of establishments from the list. Then, depending on the number that the computer selected in each specific zone, a skip rule

was defined to select truckers to survey in that zone. The skip rule for each zone was sent back to the field team.

In Cameroon, enumerators were sent to each zone with instructions as to how to apply the skip rule defined for that zone as well as how to select replacements in the event of a refusal or other cause of nonparticipation (see table A3.1 below).

### **Data Quality Control**

A management policy for the interviewers was established, and the following procedures were applied during the execution of the survey:

- (A) Daily meetings with each interviewer at the end of the day for the first revision of their questionnaire(s). The objectives of this first pass through the questionnaire were
- (1) To verify that all of the questions had been answered and that basic constraints had been respected. If the questionnaire failed this aspect of the review badly, it was returned to the enumerator to complete (through a return visit if necessary). The basic checks here included

**Table A3.1 Survey Participation by Country**

<i>Country</i>	<i>Companies</i>	<i>Approached</i>	<i>Closed</i>	<i>Refused</i>	<i>Unavailable</i>	<i>Surveyed</i>	<i>Actual</i>
Burkina Faso	Trucking companies	—	—	—	—	18	<b>16</b>
	Truckers	—	—	—	—	45	<b>45</b>
Ghana	Trucking companies	35	4	6	7	18	<b>15</b>
	Truckers	110	0	30	25	55	<b>54</b>
Cameroon	Trucking companies	52	10	12	8	22	<b>14</b>
	Truckers	120	0	40	23	57	<b>57</b>
Chad	Trucking companies	34	0	8	8	18	<b>14</b>
	Truckers	135	0	50	28	57	<b>57</b>
Kenya	Trucking companies	64	1	2	1	22	<b>21</b>
	Truckers	—	—	—	—	—	<b>55</b>
Uganda	Trucking companies	47	4	8	14	21	<b>17</b>
	Truckers	100	0	20	23	57	<b>57</b>
Zambia	Trucking companies	50	4	20	1	1	<b>19</b>
	Truckers	—	—	—	—	—	<b>45</b>

*Source:* Trucking survey.

*Note:* No available data for Burkina Faso.

- (a) that no fields had been left empty (other than explicitly skipped fields),
  - (b) that no fields had been filled in ambiguously (that is, with a dash, slash, or squiggle),
  - (c) that coded responses of DK (don't know), NA (not applicable), and R (refused to answer) seemed plausible for the field in question,
  - (d) that percentages, where required, added up to 100 percent, and
  - (e) that fields with known relationships to adjacent fields respected those relationships (for example, the year of manufacturing a vehicle is smaller than the year of purchasing that vehicle).
- (2) To take the opportunity to reinforce the enumerators' awareness of the logical links between questions, by quickly checking the more obvious ones, demonstrating what they are doing, highlighting any inconsistencies, and asking for explanations. Some of the consistency checks that might be done quickly at this stage included:
- (a) if the establishment did or did not perform cross-border operations (that is, yes or no but not NA), in which case the main point of exit that the establishment used should be listed;
  - (b) if the establishment owned lightweight vehicles, it should have the total cost of maintenance and the average fuel consumption for this type of vehicle;
  - (c) if there are some trips with overload fines in certain routes they should have the amount of overload fines per trip.
- (3) To ensure that the full consistency checks were carried out using an SPSS script after the data had been entered, and reviewed in a second meeting with the enumerator (see C below).
- (4) To assess the interviewer's ability to correctly fill out the questionnaire and to clarify any concerns regarding his understanding of the questionnaire, if necessary.
- (B) Following this review, the questionnaire was retained by the firm survey for data entry and the administration of coherence tests, unless it badly failed the first basic tests listed in A(1) above. Within a short time frame (one or two days after the first meeting), the questionnaire was entered, and the coherence and completion tests for each questionnaire were executed.
- (C) There was a second meeting with the interviewer in order to go over the results of the coherence and completion tests and, if necessary, return the questionnaire for further completion/verification.

(D) Data entry and consistency checks

- (1) When data entry was finished for the day, for each type of questionnaire for which additional cases were entered or existing cases were updated, that data files were exported to SPSS format.
- (2) The resulting SPSS script was run to open the data in SPSS.
- (3) The consistency and completion tests scripts were run in order to generate data regarding the completion status of each case and to validate the consistency checks. This procedure generated a report detailing these results as well as the completion status of the whole sample with respect to sales.
- (4) Whenever possible, this report was printed and reviewed with the enumerators and possible return visits were executed if required shortly thereafter.

(E) Completion tests

A questionnaire was considered “final” when it contained answers to 85 percent of the questions in each section. In addition, across the entire set of completed questionnaires, each variable was submitted to an 85 percent completion test. Finally, all information pertaining to the screener portion of the questionnaire had to be completed.

## **Sample Size**

The number of vehicles surveyed and details about those vehicles are shown in table A3.2.

**Table A3.2 Number of Vehicles Surveyed by Country**

<i>Country</i>	<i>Companies</i>	<i>Light weight (&lt;5 tons)</i>	<i>Medium weight (&gt;=5-7 tons)</i>	<i>Heavy weight (&gt; 7 tons)</i>	<i>Total</i>
Burkina Faso	Trucking companies	4	22	216	<b>242</b>
	Truckers	3	17	117	<b>137</b>
Ghana	Trucking companies	0	1	298	<b>299</b>
	Truckers	2	12	77	<b>91</b>
Cameroon	Trucking companies	15	10	389	<b>414</b>
	Truckers	14	32	42	<b>88</b>
Chad	Trucking companies	0	0	290	<b>290</b>
	Truckers	0	1	66	<b>67</b>
Kenya	Trucking companies	20	67	1,096	<b>1,183</b>
	Truckers	0	15	122	<b>137</b>
Uganda	Trucking companies	104	66	240	<b>410</b>
	Truckers	2	16	72	<b>90</b>
Zambia	Trucking companies	66	109	495	<b>670</b>
	Truckers	32	32	29	<b>93</b>

*Source:* Trucking survey.

*Note:* Each weight category includes trucks, trailers, and semitrailers. Tractors and towing vehicles are not included.