Cambodia - Cambodia Socio-Economic Survey 2009

National Institute of Statistics

Report generated on: April 14, 2021

Visit our data catalog at:

Overview

Identification

ID NUMBER
KHM_2009_CSES_v01_EN_M_v01_A_OCS

Overview

ABSTRACT
The CSES is a household survey with questions to households and the household members. In the household questionnaire there are a number of modules with questions relating to the living conditions, e.g. housing conditions, education, health, expenditure/income and labour force. It is designed to provide information on social and economic conditions of households for policy studies on poverty, household production and final consumption for the National Accounts and weights for the CPI.

The main objective of the survey is to collect statistical information about living standards of the population and the extent of poverty. Essential areas as household production and cash income, household level and structure of consumption including poverty and nutrition, education and access to schooling, health and access to medical care, transport and communication, housing and amenities and family and social relations. For recording expenditure, consumption and income the Diary Method was applied for the first time. The survey also included a Time Use Form detailing activities of household members during a 24-hour period.

Another main objective of the survey is also to collect accurate statistical information about living standards of the population and the extent of poverty as an essential instrument to assist the government in diagnosing the problems and designing effective policies for reducing poverty, and in evaluating the progress of poverty reduction which are the main priorities in the "Rectangular Strategy" of the Royal Government of Cambodia.

KIND OF DATA
Sample survey data [ssd]

UNITS OF ANALYSIS
Households

Scope

NOTES
Poverty reduction is a major commitment by the Royal Government of Cambodia. Accurate statistical information about the living standards of the population and the extent of poverty is an essential instrument to assist the Government in diagnosing the problems, in designing effective policies for reducing poverty and in monitoring and evaluating the progress of poverty reduction. The Millennium Development Goals (MDG) has been adopted by the Royal Government of Cambodia and a National Strategic Development Plan (NSDP) has been developed. The MDGs are also incorporated into the "Rectangular Strategy of Cambodia".

Cambodia is still a predominantly rural and agricultural society. The vast majority of the population get their subsistence in households as self-employed in agriculture. The level of living is determined by the household's command over labour and resources for own-production in terms of land and livestock for agricultural activities, equipments and tools for fishing, forestry and construction activities and income-earning activities in the informal and formal sector. Data to calculate household production were obtained from the household questionnaire and the diaries as well as data from the labour force module.

Briefly the four earlier CSES rounds have all made it possible to report sets of indicators on 8 main areas of social concern:

1. Demographic characteristics

2. Housing

3. Agriculture

4. Education

5. Labour Force

6. Health and Nutrition

7. Victimization

8. Household Income and Consumption

These 8 areas were also covered by corresponding modules in the CSES 2009, together with a diary method as well as a recall method, the other following the module design and variable content of previous rounds of the CSES with needed modifications and complements.

TOPICS

Topic	Vocabulary	URI
Demographic characteristics	CESSDA	http://www.nesstar.org/rdf/common
Housing [10.1]	CESSDA	http://www.nesstar.org/rdf/common
Agriculture	CESSDA	http://www.nesstar.org/rdf/common
Education [6]	CESSDA	http://www.nesstar.org/rdf/common
Labour Force	CESSDA	http://www.nesstar.org/rdf/common
Health and Nutrition [8]	CESSDA	http://www.nesstar.org/rdf/common
Victimization [8]	CESSDA	http://www.nesstar.org/rdf/common
Household Income and Consumption [1.1]	CESSDA	http://www.nesstar.org/rdf/common

KEYWORDS
Household

Coverage

GEOGRAPHIC COVERAGE
National Coverage

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
National Institute of Statistics	Ministry of Planning

FUNDING

Name	Abbreviation	Role
Swedish International Development Agency	SIDA	Funding

OTHER ACKNOWLEDGEMENTS

Name	Affiliation	Role
Statistics Sweden	SCB	Technical Assistance

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
CHUM PUTHIVAN	CPTV	NIS (Department)	Archivist
Saint Lundy	SLD	NIS (Department)	Archivist
Office of Chief Statistician	OCS	Food and Agriculture Organization	Metadata adapted for FAM

DDI DOCUMENT VERSION
KHM_2009_CSES_v01_EN_M_v01_A_OCS_v01

DDI DOCUMENT ID
DDI_KHM_2009_CSES_v01_EN_M_v01_A_OCS_FAO

Sampling

Sampling Procedure

In this section the sampling design and the sample selection for CSES 2009, is described. The sampling design for the 2009 survey is the same as that used for the CSES 2004. The sampling design for the 2004 CSES is described in for instance National Institute of Statistics (2005a).

The sampling frame for the 2009 survey is based on preliminary data from the General Population Census conducted in 2008. The sample is selected as a three stage cluster sample with villages in the first stage, enumeration areas in the second stage and households in the third.

The Sampling Frame

Preliminary data from the General Population Census 2008 was used to construct the sampling frame for the first stage sampling, i.e. sampling of villages. All villages except 'special settlements' were included in the frame. In all, the first stage sampling frame of villages consisted of 14,073 villages, see Appendix 1. Compared to previous years the frame used for the 2009 survey based on the census 2008 was more up to date than in previous surveys which were based on the population census 1998.

The following variables were used from the census; Province code, province name, district code, district name, commune code, commune name, village code, village name, urban-rural classification of villages, the number of households per village and, the number of enumeration areas in the village.

In the second-stage Enumeration Areas (EA) are selected in each selected village. In most villages only one EA was selected but in some large villages more than one was selected.

For the third stage, the sampling of households, a frame was constructed in field. For selected EAs the census map of the village, including EAs and residences, was given to enumerator who updated the map and listed the households in the selected EA. A sample of households was then selected from the list.

Stratification

The sampling frame of villages was stratified by province and urban and rural. There are 24 provinces and each village is classified as either urban or rural which means that in total we have 48 strata, see Appendix 1. Each stratum of villages was sorted by district, commune and village code.

Sampling

The sampling design in the CSES 2009 survey is a three-stage design. In stage one a sample of villages is selected, in stage two an Enumeration Area (EA) is selected from each village selected in stage one, and in stage three a sample of households is selected from each EA selected in stage two. The sampling designs used in the three stages were:

Stage 1. A systematic pps sample of villages, Primary Sampling Units (PSUs) was selected from each stratum, i.e. without replacement systematic sampling with probabilities proportional to size. The size measure used was the number of households in the village according to the sampling frame.

Stage 2. One EA was selected by Simple Random Sampling (SRS), in each village selected in stage 1. As mentioned above, in a few large villages more than one EA was selected.

Stage 3. In each selected EA a sample of households was selected by systematic sampling.

The selection of villages and EAs were done at NIS while the selection of households in stage three was done in field. As mentioned in section 1.1 all households in selected EAs were listed by the enumerator. The sample of households was then selected from the list.

Sample sizes and allocation

The sample size of PSUs, were, as in the 2004 survey, 720 villages (or EAs). In urban villages 10 households were selected and in rural 20 households. In all 12,000 households were selected.

Urban and rural villages were treated separately in the allocation. The allocation was done in two steps. First the sample sizes for urban and rural villages in the frame were determined and then sample sizes for the provinces within urban and rural areas were determined, i.e. the strata sample sizes.

The total sample size was divided into to two, one sample size for urban villages and the other for rural villages. The calculation of the sample sizes for urban and rural areas were done using the proportion of consumption in the two parts of the population. Data on consumption from the CSES 2007 survey was used. The resulting sample sizes for urban villages was 240 and for rural 480. (Some adjustments of the calculated sample sizes were done, resulting in the numbers 240 and 480).

Allocation of the total sample size on the strata within urban and rural areas respectively, was done in the following way. The sample size, i.e. the number of PSUs, villages, selected from stratum h, is proportional to the number of households in stratum h, i.e.

$$n(lh)=n1(Mh/\text{Sum of }Mh) \quad (1.1)$$

where,
is the sample size in stratum h, i.e. the number villages selected in stratum h,

is the total sample size of villages for urban or rural villages,

H is the number of strata in urban or rural areas,

is the number of households in stratum h according to the frame.

As mentioned above, the sample size calculations are done separately for urban and rural villages, i.e. for strata with urban villages (1.1) is used with nl = 240 and is the number of households in urban villages in the frame and for rural villages (1.1) is used with nl = 480 and is the
number of households in rural villages in the frame.

Monthly samples
In section 1.3 the selection of the annual sample was described. The annual sample was divided into 12 monthly samples of equal sizes. The monthly samples consisted of 20 urban and 40 rural villages. The division of the annual sample into monthly samples was done so that as far as possible each province would be represented in each monthly sample. Since the sample size of villages in some provinces is smaller than 12, all provinces were not included in all monthly samples. Also, the outline of the fieldwork with teams of 4 enumerators and one supervisor puts constraints on how to divide the annual sample into monthly samples. The supervisors must travel between the villages in a team and therefore the geographical distance between the villages surveyed by a team cannot be too large.

Estimation
Totals, ratios such as means or proportions were estimated for the population or for subgroups of population, i.e. domains of study. The domains were defined by e.g. region or sex. Means and proportions were estimated by first estimating totals and then calculating the ratio of two estimated totals. To estimate totals from a sample survey weights are needed.

Response Rate

The CSES 2009 enjoyed almost a 100 percent response rate. The high response rate together with close and systematic fieldwork supervision by the core group members were a major contribution for achieving high quality survey results.

Weighting

The weights are determined by the sampling design, design weights, and adjusted for nonresponse and other deficiencies such as under coverage and, to improve the precision of the estimates.
The design weight for household k in the selected enumeration area hij is given by

$W.hijk=(M.h/n.h*M.hi)*(E.hi/l)*(M.hij/m.hij)$ (2.1)

Please see the report chapter 11 (Sampling Design)

where,

Mh is the number of households in stratum h according to the frame,
Mhi is the number of households in village hi according to the frame,
nh is the number of villages selected from stratum h,
Ehi is the number of enumeration areas in village hi,
Mhji is the number of households in EA hij according to the listing of households by the enumerator and,
mhji is the number of households interviewed in EA hij

Some of the EAs have boundaries that are difficult to identify in field. In such cases there is a risk that the enumerator wrongly includes households outside the EA or excludes households within the EA. In some very large EAs the enumerator did not list all households so the number of households in the selected EA was unknown. To avoid these problems the weights were instead calculated by

$$W.hijk=(M.h/n.h*M.hi)*(M.hi/m.hij) \quad (2.2)$$

Please see the report chapter 11 (Sampling Design)

where Mhi* is the number of households in village hi according to the village chairman. Note that, in villages with only one EA the right-hand sides of (2.1) and (2.2) coincide. In villages where the EAs are of approximately the same size the right-hand sides of (2.1) and (2.2) are approximately equal. The same adjustment of the design weights were done in the 2004 survey, see National Institute of Statistics (2005a).

The weights calculated by (2.2) lead to underestimated population size. The sampling frame was constructed from a preliminary version of the population census 2008 and, the number of households per village in the frame differs from the final version. The weights were therefore adjusted using population projections and the census 2008. The population projections available at the time when the weights were calculated were preliminary and only available for age groups and sex. Therefore, also information from the census, such as population per province and household size was used to adjust the weights. Using the resulting adjusted weights, the population size is estimated to 13,966,718, and the number of households is estimated to 2,938,650.

Questionnaires

Overview

Four different questionnaires or forms were used in the survey:
1. Household listing form
The Household listing and mapping were done prior to the sampling. During the household listing the enumerator recorded household information on e.g. location, number of members and principal economic activity.

2. Village questionnaire
The Village questionnaire was used to gather basic common information on demographic information, economy and infrastructure, rainfall and natural disasters, education, health, retail prices, employment and wages, access to common prices of resources, sale of agricultural land, and recruitment of children for work.

3. Household questionnaire
The following modules were included in the Household questionnaire:
01. Initial visit
02. Education & Literacy
03. Information on migration (includes past and current migration)
04. Housing
05. Household economic activities
06. Household liabilities
07. Household income from other sources
08. Construction activities
09. Durable goods
10. Maternal health (Last pregnancy and delivery)
11. Child health (youngest child and all children under 2)
12. Health check of children under 5
13. Health care seeking and expenditure
14. Disability
15. Current economic activities (activity status during the past seven days)
16. Usual economic activity (activity in the past 12 months)
17. Victimization
18. Summary of presence in the household

4. The Diary sheet (diary method)
a. Diary for expenditure & consumption of own-production
b. Diary for household income & receipts

Minor changes were done in "kind of income" and "purpose of expenditure".

Data Collection

Data Collection Dates

Start	**End**	**Cycle**
2009-01-01	2009-12-01	N/A

Data Collection Mode

Face-to-face [f2f]

Data Collection Notes

Enumerator and supervisor training

Prior to the start of the fieldwork intensive interviewer and supervisor training were carried out. The 200 interviewers and 50 supervisors recruited were split into two groups, each group consisting of 100 interviewers and 25 supervisors. The two groups alternated so that the first group did their fieldwork during odd survey months (i.e. January, March, May, July, September, and November 2009) while the second group covered the even survey months (i.e. February, April, June, August, October, and December 2009). The training was designed with this in mind. The first group was trained in December 2008 while the second group was trained in January 2009 using premises at the NIS head office. Training of the first and second group was provided in Khmer by the appointed NIS core group and was assisted by Sida consultants. The supervisors and interviewers were jointly trained for two weeks over the 4 forms of questionnaires. During the training a special session on Gender issues relating to data collection was provided by Ministry of Women's Affair (supported by UNDP). Yet another session was held by the Cambodian Disabled People's Organization to get the enumerators better understanding the concept and definitions of disability. The Working Group on Water and Sanitation provided useful training material on the definition on improved water sources and sanitation. Training manuals are extensive and are not attached. They can however be obtained at NIS.

Interviewers and supervisors were initially divided into teams consisting of five persons (one supervisor and four interviewers), making in total 50 teams for the fieldwork. Each month 25 teams were working in the field with a workload of 10 households per interviewer. In urban areas four PSU's ("villages") were allocated to one team while in rural areas two PSU's were allocated. The fieldwork plan was designed in order to gather information from about 40 households monthly per team. For a given month the team arrived in the village three days before the first day of the interview month to tend to preparatory tasks like discussing with village authorities, filling in the Household Listing Form and thereafter sample those households to be interviewed. The Village Form was filled in by the supervisor. The Household Questionnaire had 17 sections that were filled in by the interviewer during the first visit to the household, and in the following four weeks according to the following scheme:
During a survey month different questions were asked in different weeks according to the following:
Week 1. Questions about education, migration, and housing
Week 2. Questions about economic activity, agricultural and non-agricultural business, household liabilities and other household incomes.
Week 3. Questions about construction, durable goods, health (maternal, child, general and disability)
Week 4. Questions about current economic activities, usual economic activities and Victimization
When the month ended, the team went back to the NIS headquarter in Phnom Penh.

Questionnaires from the same PSU were delivered to the NIS team for editing and coding by the supervisor in a packet including all the documents used and produced in the fieldwork, such as maps, enumeration lists and questionnaires. Appendix 6 (in the report) contains an example (the first survey month) from the allocation of teams to PSU's.
Before going to the villages the teams were briefed and introduced to minor adjustments of the interviewing procedure that were made as a result of monitoring activities and feed-back from the data processing.

Questionnaires

Four different questionnaires or forms were used in the survey:
1. Household listing form
The Household listing and mapping were done prior to the sampling. During the household listing the enumerator recorded household information on e.g. location, number of members and principal economic activity.

2. Village questionnaire
The Village questionnaire was used to gather basic common information on demographic information, economy and infrastructure, rainfall and natural disasters, education, health, retail prices, employment and wages, access to common prices of resources, sale of agricultural land, and recruitment of children for work.

3. Household questionnaire
The following modules were included in the Household questionnaire:
01. Initial visit
02. Education & Literacy
03. Information on migration (includes past and current migration)
04. Housing
05. Household economic activities
06. Household liabilities
07. Household income from other sources
08. Construction activities
09. Durable goods
10. Maternal health (Last pregnancy and delivery)
11. Child health (youngest child and all children under 2)
12. Health check of children under 5
13. Health care seeking and expenditure
14. Disability
15. Current economic activities (activity status during the past seven days)
16. Usual economic activity (activity in the past 12 months)
17. Victimization
18. Summary of presence in the household

4. The Diary sheet (diary method)
a. Diary for expenditure & consumption of own-production
b. Diary for household income & receipts

Minor changes were done in "kind of income" and "purpose of expenditure".

Data Collectors

Name	Abbreviation	Affiliation
National Institute of Statistics	NIS	Ministry of Planning

Data Processing

Data Editing

Data editing and coding:
The NIS team commenced their work of checking and coding in beginning of February after the first month of fieldwork was completed. Supervisors from the field delivered questionnaires to NIS. Sida project experts and NIS Survey Manager helped solving relevant matters that became apparent when reviewing questionnaires on delivery.

Basic instructions
All questionnaires from each PSU were delivered to editors and coders by supervisor. The editors and coders were responsible for handling the questionnaires from the brought from the field supervisor's until finishing the process of checking and coding. When checking and coding a red pen was used in the questionnaire.

How the workflow is organised at the office:
Data editing and coding is an important part of the overall data processing for CSES. In brief, the implementation of data editing and coding comprise the following functions:
When a field supervisor delivered questionnaires from a PSU the delivery contained a set of mappings, listings, village questionnaires, household questionnaires and diary forms. Editors and coders started checking each PSU including mapping information and all other forms. Field supervisor had to wait for editor and coder's checking. If any problem occurred, editor had to immediately ask field supervisor to correct the error.
After corrections were completed, editor started the coding process. The code to be used included e.g. crop-code, occupation, industry code, income and expenditure code, and unit code. When editor encountered a mistake which could not be corrected directly by editor it had to be discussed with the supervisor or called back to enumerator.
After checking and coding was finished, the data editor staff put all documents from the PSU into a designated box labelled with the PSU number and sent it to the data-entry operator.
In case the data-entry operator encountered any mistakes caused by checking and coding, the operator sent the questionnaire back for re-edit and checking.
Editing and coding proceeds every month and is done one week before data entry starts.

Other Processing

Training: In December 2008, the data processing team participated in a training course for enumerators and supervisors. The main objective of the training was to identify anomalies in the questionnaire and also discuss certain ideas raised during training sessions to avoid and reduce future mistakes. From January 2009 and onwards, the supervisor for data editing and coding took part in reviewing problems raised by instructors and enumerators encountered during fieldwork interviews.

In late 2006 and beginning of 2007 a new system for data processing and storage were introduced for the Cambodia Socio Economic Survey (CSES). It includes a relational database system for storing CSES data in SQL format and an application framework developed in-house for data-entry. Since NIS staff already was familiar with Visual Basic and Microsoft SQL Server database software the transition from previous data processing system was feasible. A modern network infrastructure within the NIS was also implemented to host the new CSES system and facilitate for concurrent data-entry.
The application and storage platform developed in 2006 and supervised by Statistics Sweden consultancy has since been used consecutively for all CSES data processing from 2007 and onwards.
The database contains data tables for all modules comprising the CSES household, village and diary questionnaires. There are also code-tables used for data integrity controls during data-entry and tables for data management including error lists. In all the database counts a total of 185 tables divided by:
Table Count
Data tables 39
Code tables 129
Management tables 17

To facilitate for easier data retrieval there are also a set of views or virtual tables available in the database.
Data in the system is for the most part processed by three distinct application components all developed in Microsoft Visual Basic 6.0.
The CSES editing component: is used for entering household information from cover page of questionnaire such as; PSU number, household number and number of members in household.
The CSES entry component: is used for entering data from each CSES module. This is the main component for data

processing.
The CSES management component: is used to correct errors and view information about operator statistics.

All database modules as well as application components have since 2008 been maintained and improved by staff from the NIS ICT department.

Work flow of CSES data processing system
Step 1 Questionnaires sent from field operators arrives monthly at NIS and is taken care of by data processing staff
Step 2 Questionnaires are updated with appropriate codes for household identification. They are checked and edited for any apparent errors or misunderstandings from the field operator. All changes are written to the questionnaire.
Step 3 Data about number of rows for each module are entered into the management module. As well as number of households per PSU. These practices are to ensure that all data rows are entered.
Step 4 Data-entry of all modules including Household, Diary and Village.
Step 5 After finished data entry an iterative error correction phase is started and run from the database server. Any errors from data controls are visible in the management module.

Data Appraisal

Estimates of Sampling Error

In order to provide a basis for assessing the reliability or precision of CSES estimates, the estimation of the magnitude of sampling error in the survey data shall be computed. Since most of the estimates from the survey are in the form of weighted ratios, thus variances for ratio estimates will thus be presented.

The Coefficients of Variation (CV) on national level estimates are generally below 4 percent. The exception is the CV for total value of assets where there are rather high CVs especially in the urban areas, which should be expected. The CVs are somewhat higher in the urban and rural domains but still generally below 7 percent. For the five zones, the average CVs are in the range 5 to 13 percent with a few exceptions where the CVs are above 20 percent. For provinces the CVs for food consumption are 9 percent on average.

The sample take within Primary Sampling Units (PSU) was set to 10 households per PSU in the CSES 1999. When data on variances became available, it was possible to make crude calculations of the optimal sample take within PSU. Calculations on some of the central estimates in the CSES 1999 show that the design effects in most cases are in the range 1 to 5.

Intra-cluster correlation coefficients have been calculated based on the design effects. These correlation coefficients are somewhat high. The reason is that the characteristics that are measured tend to be concentrated (clustered) within the PSUs. The optimal sample size within PSUs under different assumptions on cost ratios and intra-cluster correlation coefficients was then calculated. The cost ratio is the average cost for adding a village to the sample divided by the average cost of including an extra household in the sample. In the CSES, it was chosen to adopt a fairly low cost ratio due to the fact that the interview time per household is long. Under this assumption the optimal sample size is probably around 10 households per village for many of the CSES indicators.