D R A F T

**Employment Survey in Palm Oil Plantations in Malaysia:**

**Sample and Questionnaire Design**

Farhad Mehran, ILO Consultant

[mehranxfarhad@yahoo.com](mailto:mehranxfarhad@yahoo.com)

and

Federico Blanco, ILO Fundamentals

[blanco@ilo.org](mailto:blanco@ilo.org)

10 May 2018

Contents

1. Introduction
2. Sample design

* Basic approach
* Sample size determination
* Direct sampling of plantations
* Indirect sample of enumeration blocks (EBs)
* Sample selection of workers within EBs

1. Questionnaire design
   * + Employment module
     + Forced labour module
     + Child labour module
2. Follow-up activities

* Pilot survey
* Main survey

Annexes

1. Excel file: Sample size calculation
2. Excel file: Sampling MPOB Estates
3. Excel file: Sampling Independent Smallholders
4. **Introduction**

Malaysia is the second largest producer of palm oil in the world after Indonesia. Total production of palm oil in Malaysia was 19,962,000 tonnes in 2015 on about 5,600,000 hectares of land, representing about 32 percent of total world production in that year.[[1]](#footnote-1) Most of the production is exported (17,454,000 tonnes) and Malaysia is also the second largest exporter of palm oil after Indonesia, making up about 37 percent of world’s export of palm oil in 2015. Malaysian plantations are considered to be the most productive globally, yielding on average of 18.48 fresh fruit bunches (FFB) of palm oil per hectare of planted land, against an average of 7 FFB per hectare in countries other than Malaysia and Indonesia.

The palm oil industry is labour intensive, especially in the palm oil plantations, in activities such as harvesting, collecting fruits, weeding and general work. A census of plantations carried out in Malaysia in 2010 estimated that the industry engaged some 446,368 workers in 2010, mostly foreign workers (69 percent) composed mainly of Indonesians but also of other nationalities, in particular from Bangladesh, Thailand and Myanmar. The “List of Goods Produced by Child Labor or Forced Labor” issued in 30 September 2016 by the US Department of Labor identified Oil (palm) in Malaysia as a good produced by forced labour and child labour.[[2]](#footnote-2)

To estimate the extent of the phenomena and examine the employment situation of workers in the palm oil sector in Malaysia, the government in collaboration with the ILO Branch on Fundamental Principles and Rights at Work has agreed on a project to conduct a nationally representative survey in the Malaysian palm oil sector in 2018. An initial sample design was proposed and discussed during a video conference with experts from the ILO and from the Department of Statistics Malaysia (DOSM), Ministry of Plantation Industries and Commodities (MPIC) and Ministry of Human Resources (MOHR). The proposed sample design was further discussed with experts from the Malaysian Bureau of Labour Statistics (MBLS) at the occasion of the meeting of the ILO Working Group on Statistics of Forced Labour in Rio de Janeiro, 20-22 March 2018. An ILO mission was then fielded from 16 to 20 April 2018 to review and finalize the sample design with the staff of DOSM and discuss the elements of a pilot survey to be conducted in May 2018.

The present document reports on the outcomes of the ILO mission in April 2018. It is organized in four sections. Following this introductory section, Section 2 describes in details the agreed sample design of the survey. Section 3 lays out the basic elements of the questionnaire to be finalized in a subsequent mission in mid-May by the ILO expert, Michaelle de Cock. The final section, Section 4, reports on the agreed pilot survey and the follow-up activities in preparation of the implementation of the survey. The calculations of the sample size and the control of the parameters are given in an Excel file (Sample size.xlsx) considered as Annex A of the present document. The sample selection for large estates is given as Annex B in a separate Excel file (LIST OF MPOB ESTATES BY CATEGORIES 5.xls) and that for smallholders as Annex C in Excel file (Smallholders Sample Selection 20 April 2018 for farhad.xlsx).

1. **Sample design**

The main challenge of the sample design is to draw a probability sample of palm oil plantation workers at their living quarters rather than at their place of work. Experience has shown than in household-based surveys reaching the workers at their living quarters, the respondents are likely to feel freer to talk about their work experience than they would at their workplace in the presence of their employer or work colleagues. Because no sample frame of plantation workers exists, the sample is selected indirectly on the basis of area units (Enumeration Blocks) linked to palm oil plantations. The procedure is explained in detail below.

* **Basic approach**

The basic sampling approach may be described as multi-frame indirect sampling of workers at their living quarters. It is indirect sampling because no direct sampling frame of living quarters of palm oil plantation workers exists, and sampling has to be carried out indirectly through the initial selection of palm oil plantations. It is multiple frame because of the distinct list of large private and government estates from that of independent smallholders with possible overlap of the two lists. A schematic representation of the sampling approach is shown in the following diagram.



Starting from the list of large private and government estates provided by the Malaysian Palm Oil Board (MPOB) containing 5097 plantations, a stratified sample of plantations is drawn with probabilities proportional to size defined in terms of hectares. TheThe sample plantations are then linked to the area frame of the Department of Statistics Malaysia (DOSM). Each unit of the frame is an Enumeration Block (EB) covering on average about 80 to 120 living quarters. It should be noted that a sample plantation may be linked to one or more EBs. Similarly, each EB may be linked to one or more plantations.

In order to increase the likelihood of covering workers living outside the plantation and workers engaged in a plantation other the sample plantation, each linked EB is paired to another neighboring EB, called paired EB. The paired EBs are selected clockwise, one for each linked EB. In the final stage of sampling, the living quarters in each linked or paired EB is screened for living quarters with plantation workers. In linked EBs, up do 20 plantation workers are sampled according to an adaptive cluster sampling scheme described later in this section. In paired EBs, the sample-take is set lower, at 10 plantation workers.

In the case of the independent smallholders, the sampling frame is much larger covering 292904 plantations. The sampling frame is first stratified by geographic region (Negeri) and sorted by size of plantation. Then, a sample of plantations is drawn by systematic sampling with equal probabilities from the sorted file. As in the case of the estates frame, the sample of smallholder plantations are linked to the area frame of EBs, but in this case the linked EBs are not paired with other neighboring EBs. This is because smallholder plantations are likely to be within the boundaries of the enumeration blocks to which they belong and therefore pairing them with neighboring EBs are less likely to provide additional information on smallholder employment characteristics. For the same reason, the sample of workers to be drawn from the linked EBs of the smallholder frame is set at a lower number (10 workers per linked EB) as compared with that of the estate frame (20 workers per linked EB). These parameters may be altered as required in the Excel file on sample size determination.

The sample selection of workers within EBs, linked or paired, proceeds as follows. First, starting from a few starting points (core addresses) and following a random route, the living quarters of the sample EBs are screened for identification of households with members working or helping in palm oil plantations including children and adolescents 5 to 17 years old. A positively screened household falls in the sample and is interviewed. The interviewer then proceeds to the two immediate neighbors for further screening and possible interviewing. A negatively screened household falls outside the scope of the sample and it is not interviewed. In this case, the interviewer does not go to the neighboring living quarters, but proceeds the normal random route for further screening and interviewing. The methodology provides for a stopping rule when the number of sample workers reaches the target sample-take of 20 or 10 workers per EB as the case may be. The methodology also provides to dropping rule when the screening process takes too long without reaching a positively screened household.

* **Sample size determination**

The sample size of plantations to be drawn from each of the two frames is determined backwards, first the required sample size of workers for a given precision requirement, then the corresponding number of EBs given the sample-take per EB, and then the required number of plantations given the average number of links between EBs and plantations.

* Sample of size of workers

The required sample size of workers is determined on the basis of the conventional expression for sample size determination,

where p is the prescribed value of the prevalence of the target variable, here forced labour, q is the complement value, 1-p, *deff* is the design effect, ε is the specified margin of error of the estimate, r is the assumed response rate and γ is the assumed rate of non-matching plantations and EBs.

The value of p is determined on the basis of the world average prevalence rate of forced labour, 0.006 per person, converted to the prevalence per worker, 0.014 per worker, using data on the labour force participation rate in Malaysia and the share of working age population, 15-64 years old, to total population:

The design effect is set at a mid-value, *deff*=1.5, corresponding to what is used in many national labour force surveys. The design effect is an adjustment to the sample size to take into account the departure of the sample design from a simple random sampling scheme. The margin of error is set at ε=0.003. It corresponds to a confidence interval of estimation of the prevalence of forced labour within about one-fifth of the prescribed value, that is 0.014±0.003, at α=5 percent level of significance, where 1.96, the tail value of the standard normal distribution.

Finally, the sample size is adjusted to compensate for non-response among workers and for sample loss due to failures in matching sample plantations and Enumeration Blocks. The rate of non-response among workers is considered to be around 5 percent corresponding to the typical value obtained in the regular labour force survey conducted by DOSM. This gives the response rate of r = 0.95 used in the expression of the sample size determination. For matching failures, a much higher rate is set, γ = 0.25, as preliminary experience has shown that many plantations cannot be matched with the Enumeration Blocks because the addresses in the frame are either incorrect or incomplete, or as in many cases referring to the address of the owner or manager rather than the address of the plantation itself.

The resulting sample size is n=12028 workers per domain. The sample design considers three domains: Peninsula Malaysia, Sabah and Sarawak. Peninsula Malaysia excludes the urban states, W.P. Kuala Lumpur, W.P. Labuan and W.P. Putrajaya, where there are virtually no palm oil plantations. The total design sample size is thus n=36084 workers. Table 2 gives the allocation of the sample in each domain in terms of the two sampling frames, government and private plantation estates (Frame I) and independent smallholders (Frame II).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. **Sample size in terms of number of workers** | | | | |
| Domain | | Government & private estate plantations  (Frame I) | Independent smallholders  (Frame II) | Total |
| A | Peninsula Malaysia | 8146 | 3882 | 12028 |
| B | Sabah | 8685 | 3343 | 12028 |
| C | Sarawak | 8729 | 3299 | 12028 |
|  | **Total** | **25560** | **10524** | **36084** |
|  | Sampling rate | 4.2% | 9.5% | 5.1% |

The sample design provides for oversampling of workers in independent smallholders on the assumption that the prevalence of forced labour is likely to be higher among them than among workers in government and private estate plantations. Oversampling is achieved by allocating the sample in each domain in the ratio of the square-root of the number of workers or equivalently the number of hectares covered by plantations of the different frames. As shown in the last line of Table 2, the sampling rate of workers of the independent smallholders indicates is 9.5 percent compared with the sampling rate of 4.2 percent for workers in government and private estate plantations.

* Sample size of EBs

The conversion of the sample size requirement in terms of number of enumeration blocks (EBs) may be obtained based on the sample-take per type of EBs. As mentioned earlier, the sample-take is 20 workers per linked EB and 10 workers per paired EB according to the sample design from frame I. In the case of frame II, the sample-take is 20 workers per linked EB and no paired EB. The resulting sample size requirements in terms of the number of EBs are given in table 3 below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 1. **Sample size in terms of number enumeration blocks (EBs)** | | | | | |
| Domain | | Government & private estate plantations  (Frame I) | | Independent smallholders  (Frame II) | Total |
| Linked EBs | Paired EBs | Linked EBs |
| A | Peninsula Malaysia | 271 | 271 | 194 | 736 |
| B | Sabah | 289 | 289 | 167 | 745 |
| C | Sarawak | 290 | 290 | 165 | 745 |
|  | **Total** | **850** | **850** | **526** | **2236** |

It can be verified, for example, that for Peninsula Malaysia, the sample size from frame I in terms of EBs:

271 linked EBs x 20 workers per linked EB + 271 paired EBs x 10 workers per paired EB

= 8130 workers

is consistent with the sample size in terms of number of workers (8146) given in table 2, except for rounding errors.

* Sample size of plantations

The final step in the determination of the sample size is the conversion of the sample size in terms of number of plantations. This is calculated on the basis of the assumption that each plantation in frame I is linked on average to about 1.2 EBs. In the case of frame II, because of the small size of the plantations, it is assumed that each plantation is linked to just 1 EB. Table 4 shows the resulting sample size requirement in terms of number of plantations from each frame and each domain.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. **Sample size in terms of number of plantations** | | | | |
| Domain | | Government & private estate plantations  (Frame I) | Independent smallholders  (Frame II) | Total |
| A | Peninsula Malaysia | 226 | 194 | 420 |
| B | Sabah | 241 | 167 | 408 |
| C | Sarawak | 242 | 165 | 407 |
|  | **Total** | **710** | **526** | **1236** |
|  | Sampling rate | 13.9% | 0.2% | 0.4% |

It should be noticed that the conversion of the number of sample EBs to the number of sample plantations is calculated on the basis of linked EBs only, as paired EBs do not influence the number of plantations that should be sampled. Thus, for example, for Peninsula Malaysia, the sample size requirement from frame I is 271/1.2 =226 plantations, where 271 is the expected sample size of linked EBs obtained from table 3 and 1.2 is the average number of EBs linked to a plantation in frame I. It should also be mentioned the sample design ignores at this stage any possible overlap between frame I and frame II. This issue if it occurs will be handled at the estimation stage with the calculation of the sampling weights.

Overall, the sample design requires the sample selection of 1236 plantations, 710 plantations from frame I on government and private estates and 526 plantations from frame II on independent smallholders, corresponding to sampling rates of 13.9 percent and 0.2 percent from the two frames, respectively. In the remainder of this section, the procedure for sample selection of the plantations and the corresponding EBs are described, followed by a description of the sample selection of workers within the sample EBs.

* **Direct sampling of plantations**

For each frame, the sample size of plantations is first allocated among strata and then sample selection is carried out within strata by probability proportional to size in frame I and by equal probability in frame II. Stratification in frame I involves three variables: state, type of plantation (government or private estate), and size of plantation (below 200 hectares, 200 to 999 hectares and 1000+ hectares). The thresholds have been chosen to lead to an equal share of plantations in each size group. Overall 71 strata are thus formed for 13 states, 2 types and 3 size groups of plantations with certain type and size group of plantations not present in some states. In the case of frame II, stratification has been limited to one variable (state) as the other variables (type and size of plantation) are not relevant or effective for stratification of independent smallholders.

For both frames, the sample allocation among strata uses as a compromise between equal allocation and proportional allocation. Equal allocation would treat each stratum equally and thus large strata would be assigned the same sample size as small strata. Proportional allocation would assign a sample size proportional to size of the stratum, resulting with very few observations in small strata. The square root allocation provides a compromise and allocates the sample size among strata in proportion to the square-root of the size of stratum, measured here in terms of number of hectares, essentially the same as the number of workers. The square root allocation corresponds to the parameter α=0.5 in the Excel files of Annexes A, B and C. The method of allocation may be changed by a different specification of the parameter (α=0 for equal allocation and α=1 for proportional allocation).

The resulting sample allocation is shown in Table 5 below. In absolute terms the largest allocations are in Sabah and Sarawak. This is because these two states are also statistical domains under the survey design. In relative terms, however, the sampling rate in all states are close to each other within a narrow margin, from about 0.1 percent in Johor and Perak to about 1.3 and 1.8 percent in Kelatan and Luar Negeri, respectively. The only exception is Perlis where the sampling rate is about 7.5 percent. The total number of plantations in this state is very limited, 3 government and private estates and 37 plantations of independent smallholders. Out of this total of 40 plantations, 3 are selected in the sample corresponding to a sampling rate of 3/40=7.5%.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1. **Sample allocation of plantations among states and types of plantation** | | | | | | |
| State | | Frame I | | Frame II | Total sample | Sampling rate  (%) |
| Government estates | Private  estates | Independent smallholders |
| 01 | JOHOR | 13 | 24 | 48 | 85 | 0.1% |
| 02 | KEDAH | 4 | 12 | 13 | 29 | 0.4% |
| 03 | KELANTAN | 7 | 13 | 7 | 27 | 1.3% |
| 04 | LUAR NEGERI | 0 | 0 | 1 | 1 | 1.8% |
| 05 | MELAKA | 4 | 10 | 10 | 24 | 0.6% |
| 06 | NEGERI SEMBILAN | 9 | 14 | 12 | 35 | 0.6% |
| 07 | P. PINANG | 1 | 4 | 11 | 16 | 0.4% |
| 08 | PAHANG | 17 | 20 | 19 | 56 | 0.4% |
| 09 | PERAK | 12 | 22 | 35 | 69 | 0.1% |
| 10 | PERLIS | 2 | 0 | 1 | 3 | 7.5% |
| 11 | SABAH | 53 | 187 | 167 | 397 | 1.0% |
| 12 | SARAWAK | 45 | 190 | 165 | 400 | 0.9% |
| 13 | SELANGOR | 5 | 13 | 27 | 45 | 0.2% |
| 14 | TERENGGANU | 9 | 9 | 10 | 28 | 0.7% |
|  | **Total** | **181** | **518** | **526** | **1225** | **0.2%** |

For each frame, the probability of selection of a plantation is calculated and included as a variable in the data file. The data files are then sorted by strata and within strata by district for frame I and by strata and within strata by size of plantation for frame II. Based on the sorted files, the sample selection of plantations is then carried out using the algorithm shown in diagram 6.



A simple random number between 0 and 1 is drawn for each sampling frame. Then the probabilities are sequentially cumulated starting with the negative value of the random number. If the integer value of the cumulated probabilities changes its value, the unit is in the sample, otherwise it is not. The probabilities of selection are determined in line with the sample design, that is probability proportional to size within strata in frame I and equal probability within strata in frame II. Thus, for frame I,

where c is a factor of proportionality and xih is the measure of size of plantation I in stratum h in hectares. The parameter c is calculated by newton iteration to ensure that the probabilities are within 0 and 1. Starting with the sample allocation in the corresponding stratum h, co=nh, the value of c is obtained by

where min is the minimum function, sign is the sign function and pih=xih/Σhxih. For frame II, the calculations of the probabilities of selection are much simpler. They are given by the equal probability sampling scheme within strata,

where nh is the sample allocation in stratum h and Nh is the total number of plantations of stratum h in the sampling frame.

The calculations are programmed in the Excel files of Annexes B and C. Each activation of a file produces a new random number and generates a random sample according to the specifications of the sample design. The specifications are controlled by the parameters given in the Excel file of Annex A.

* **Indirect sample of Enumeration Blocks**

The sample enumeration blocks (EBs) are obtained indirectly on the basis of their links with the sample of plantations drawn according to the sample design described in the preceding section. Diagram 7 shows the linkage between the sampling population (i.e., the plantations listed in the rows) and the target population (i.e., the enumeration blocks listed in the columns), where N is the total number of plantations in the sampling frame and M is the total number of EBs in the census frame.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **7. Indirect sampling of Enumeration Blocks** | | | | | | | | | | | |
| Sampling population (Plantations) | Target population (Enumeration blocks, EBs) | | | | | | | | | | Total |
| 1 | 2 | . | . | . | j | . | . | . | M |
| 1 |  |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |  |
| . |  |  |  |  |  |  |  |  |  |  |  |
| . |  |  |  |  |  |  |  |  |  |  |  |
| . |  |  |  |  |  |  |  |  |  |  |  |
| i |  |  |  |  |  | θij |  |  |  |  |  |
| . |  |  |  |  |  |  |  |  |  |  |  |
| . |  |  |  |  |  |  |  |  |  |  |  |
| . |  |  |  |  |  |  |  |  |  |  |  |
| N |  |  |  |  |  |  |  |  |  |  |  |
| Total |  |  |  |  |  |  |  |  |  |  |  |

The linkage is specified as the intersection of a given row and column by

A link exists between a plantation and an enumeration block if the EB includes any part of the plantation or if the EB is a neighboring EB paired to the plantation. Thus, a plantation may be linked to more than one EB, and vice versa an EB may be linked to more than one plantation. For a given sample of plantations s = {i: 1,2,…,n}, where n is the sample size, we obtain an indirect sample of EBs, say, sEB = {j: 1,2,…,m}, where m is the number of indirect sample of EBs. It is important to note that a given EB could have been selected in the sample either as a linked or paired EB associated with a sample plantation or as a linked or paired EB associated to another plantation in the sampling frame, though not appearing in the plantation sample. This multiplicity sampling is treated in the calculation of the sampling weights. In indirect sampling with multiplicity, the sampling weight of the enumeration blocks are calculated by the weight-share method as follows. [[3]](#footnote-3)

Let j denote a sample enumeration block, its sampling weight wj is obtained by

where πi is the probability of selection of plantation I as calculated earlier and

where *θ*ij is the link function. The upper limit N in the last expression indicates that the calculation of the sampling weights involves information beyond what is contained in the sample. In order to calculate θ+j it is necessary to know the number of links the EB j has with plantations of the sampling frame, not just the one falling in the sample. Thus, it is important to keep record on the number of plantations linked to the sample EB. This information is generally obtainable from the EB map.

* **Sample selection of workers within enumeration blocks**

The final stage is the selection of workers within the sample enumeration blocks. The proposed procedure is adaptive cluster sampling along a random route. Adaptive cluster sampling refers to a sampling design in which an initial set of units is selected by some probability procedure, and whenever the variable of interest of a selected unit satisfies a given criterion, additional units in the neighbourhood of that unit are added to the sample.[[4]](#footnote-4) The procedure is efficient if palm oil workers can be assumed to be living in dwellings close to each other. A version of the methodology has been used in several European countries as part of the Second European Union Minorities and Discrimination Survey.[[5]](#footnote-5)

To fix ideas, consider an enumeration area with 25 living quarters, 6 with palm oil workers members, marked in dark colour in diagram 8. A living quarter is initially selected with equal probability, say living quarter 11, and screened for the presence of palm oil workers. As in this case, there is a palm oil worker, the interviewer is instructed to carry the interview and proceeds to the neighbouring living quarters, i.e., living quarters 10 and 12. Otherwise, the instruction would have to proceed on the random route assigned for the enumeration block, say to the next fifth living quarter. In this case, both living quarters 10 and 12 are also screened positive, so the interviewer carries the interview and moves to the next living quarters 9 and 13. In living quarter 9, there is no palm oil worker, so the interviewer stops moving in that direction. But living quarter 13 is screened positive, so the interviewer makes the interview and continues the process in that direction going to living quarter 14. In living quarter 14, no palm oil worker is found and therefore the process stops.

**8. A schematic representation of adaptive cluster sampling: An enumeration block (EB) with 25 living quarters, 6 with palm oil workers and 19 without palm oil workers**

It can be shown that the procedure is efficient. Under adaptive cluster sampling, in this example on average 2.04 living quarters are to be screened, finding 0.8 of them with palm oil workers. The success rate is therefore 0.8 to 2.04, or about 39 percent. The corresponding rate under simple random sampling is 6 to 25, or about 24 percent.

In practice, the implementation of adaptive cluster sampling involves a number of considerations. First, the set of initial living quarters. In the EU minorities and discrimination survey, their number was set at 60% of the number of addresses that would otherwise be required to deliver the same number of interviewers by direct screening. Thus, an EB with 100 living quarters, and a 5-percent rate of palm oil workers would be assigned 3 living quarters for initial sampling, rather than just one in the simplified example described above.

The second consideration is the number of neighbouring living quarters to screen in the adaptive cluster sampling scheme. In the EU survey, if a core address was found to contain eligible members of the target population, the interviewer was instructed to screen the two neighbouring addresses on either side of the core address rather than one in our simplified illustration. The expansion to 2 adjacent addresses provide for the possibility of drawing a larger set of target units into the adaptive cluster. For example, in Diagram 7, a two-adjacent address rule would draw all living quarters 7 and 8, and 10, 11, 12, and 13 into the same cluster, because having reached living quarter 8, the interviewer would in this case go to living quarter 10 as well as 9, and would thus hit again a living quarter with palm oil worker and proceed more efficiently.

The third consideration is the stopping rule. In general, the adaptive chain stops if the interviewer reaches a neighbouring living quarter that did not contain anyone who was eligible, or if eligibility could not be established, or if living quarter already screened is reached. In the present context, the stopping rule is when the chain stops or when the sample take for the enumeration block is reached, that is 20 palm oil workers in linked EBs of frame I, and 20 palm oil workers in paired EBs of frame I or linked EBs of frame II.

The fourth consideration is the dropping rule. To improve the efficiency of the field operations, it is wise to establish a dropping rule so as to avoid excessive screening in EBs with no or very few living quarters having palm oil workers. Thus, an EB would be dropped from the sample (and a replacement issued) if its initial sample outcomes did not meet a minimum threshold of fieldwork success. The threshold could be set so as to ensure that the probability of incorrectly dropping a sufficiently ample EB is below a certain percentage (10%), that is the type-1 error rate is less than 10 percent. The threshold *ndrop* may be calculated based on the negative binomial distribution,

where p is the rate of palm oil workers in the EB. Thus, if it is desired to safeguard against erroneous drop of EBs with palm oil workers rate of at least 10 percent, then the above expression leads to the threshold of ndrop=22. This means that if after 22 screening of living quarters, all did not contain any palm oil worker, then the EB could be dropped and replaced.

Finally, there is the issue of weight calculations. Adaptive cluster sampling may be regarded as a special case of indirect sampling with multiplicity. Accordingly, the sampling weight of a living quarter k in enumeration block j may be expressed as

where μk is the number of living quarters with palm oil workers in the cluster, mj is the total number of living quarters in the enumeration block j and mo is the number of initial addresses used in the adaptive cluster sampling scheme. In the numerical example of diagram 7, mj=25, mo=1,and μk=2 for living quarters k=7 and k=8, and μk=4 for living quarters k=10, k=11, k=12, and k=14. The corresponding sampling weights can then be calculated as wk(j)=25/2 for k=7 and k=8, and wk(j)=25/4 for k=10, k=11, k=12 and k=13.

In closing, it should be mentioned that in practice the calculation of the sampling weights needs to take into account non-response and cases where the outcome of the screening could not be clearly determined. Weight adjustment should also take into account the fact the full length of the adaptive chains was not always established where the fieldwork was stopped upon reaching the sample-take for that EB.

1. **Questionnaire design**

The proposed questionnaire for the survey of employment of palm oil plantations in Malaysia is based on three modules dealing with employment, forced labour and child labour. The modules are described in turn below.

* Employment module

The employment module is built around the questionnaire of the national labour force, salaries & wages survey (LFS) conducted on quarterly basis by DOSM. Apart from the cover page and end page, the LFS questionnaire contains: a page on household identification and interview date and timing (section 1); a household roster for listing the household members and recording their main socio-demographic characteristics such as sex, age, ethnic group, citizenship, marital status, educational attainment, etc. (section 2); labour force particulars of household members aged 15 years and over, including characteristics of principal and secondary jobs (section 3); salaries and wages for government and private sector employees (section 4).

For the present palm oil plantation survey, it is proposed to maintain the basic elements of the LFS questionnaire with a few modifications for adaptation to the broader scope of the survey which now covers children below the cut-off age of 15 years. The proposed modifications are as follows:

Replace question S1 with the following new wordings:

|  |
| --- |
| S1\_new. Did you work at least ONE HOUR during the (Reference Week) for pay or profit or for family gains? (Include own account work). Work include:   1. Any work on his/her own or the household’s plot, farm, food garden, or help in growing farm produce or in looking after animals for the household? 2. Online sales activities such as selling clothes, accessories, cup-cakes, etc. 3. Assist family members to operate their enterprises/businesses 4. Activities carried out at home such as folding boxes, shelling of prawns, sending school children and sewing beads 5. Yes………….………… 6. No……………………... |

Maintain questions (S2 to S6) on employment identification

Omit questions (S7 to S12) on unemployment

Maintain questions (S13 to S15) on employment characteristics at principal job

Replace questions (S16 to S19) on secondary job with the following new questions (S16\_new to S19\_new) addressed only to children 5-17 years-old:

|  |
| --- |
| S16\_new. Was the work reported in S13, S14 and S15, your main activity during the past 12 months?   1. Yes………….………… 2. No………………………………………………………………. (go to S17\_new) 3. Did not report any work on S13, S14 and S15 (go to S17\_new) |

|  |
| --- |
| S17\_new. Even though you were not engaged in an employment activity last week, did you work at least ONE HOUR during the (past 12 months) for pay or profit or for family gains? (Include own account work). Work include:   1. Any work on his/her own or the household’s plot, farm, food garden, or help in growing farm produce or in looking after animals for the household? 2. Online sales activities such as selling clothes, accessories, cup-cakes, etc. 3. Assist family members to operate their enterprises/businesses 4. Activities carried out at home such as folding boxes, shelling of prawns, sending school children and sewing beads 5. Yes, for pay or profit (i.e., employee or own-account worker) 6. Yes, for family gain (i.e., contributing family worker) 7. No……………………... |

|  |
| --- |
| S18\_new. Occupation   1. What is your occupation? 2. Can you describe your duties/nature of your job? |

|  |
| --- |
| S19\_new. Industry   1. What is the name and address of your establishment or the establishment which you are employed? 2. What are the main activities/product of this establishment? 3. In which state/country is your usual place of work? |

* Forced labour module

The forced labour module is a series of 21 questions to be administered after the employment module addressing all household members working or helping in palm oil plantations including children and adolescents 5 to 17 years old. Workers include employees, apprentices, and contributing family workers engaged in the production of palm oil, but exclude the employers, owners, and managers of the plantation. The target population may be identified on the basis of questions S15 on status in employment and questions S14 and S19 on branch of economic activity of current job, and of job held last year in the case of children 5 to 17 years old.

The purpose of the forced labour module is to measure the prevalence of forced labour among workers in the palm oil plantations and to understand its characteristics. The ILO Forced Labour Convention, 1930 (No. 29) defines forced or compulsory labour as “all work or service which is exacted from any person under the menace of any penalty and for which the said person has not offered himself voluntarily” (Article 2).[[6]](#footnote-6) For the purpose of statistical measurement, forced labour may be characterised by the two basic criteria embedded in the ILO Convention, namely:

1. “involuntary work”, i.e. engagement in an economic activity without informed consent such as unfree recruitment into a job (slavery or debt bondage); working overtime or for a longer period of time than agreed; working at substandard or with no wages, etc.
2. “work under menace of any penalty”, i.e., working under threat or menace of a threat of a penalty such as physical or mental violence, withholding of wages or other promised benefits, withholding of travel or identity documents, threat of denunciation or deportation, etc.

The concept of “work” is defined as “any activity performed by persons of any sex and age to produce goods or to provide services for use by others or for own use” in line with the latest international statistical standards.[[7]](#footnote-7) In particular circumstances, the scope of work for the measurement of forced labour may be broaden to include activities such as begging that go beyond the scope of production of goods and services of the international standards.

The forced labour module attempts to operationalize these concepts into a series of questions to be implemented under the less sensitive label of “employment conditions”. The proposed questions have been prepared by ILO staff on the basis of the experienced gained in a number of surveys conducted by the ILO in countries of different regions in the past decade or so. The question wordings, as well as their answer categories and formats, are subject to review of DOSM and should be field tested before finalization. The proposed questions are listed below. Responses to questions R2 and R3 are meant to be used for the calculation of the sampling weights. They provide information on the possibility multiplicity of sample selection of the worker. Questions R4 to R8 concern recruitment and questions R9 to R13 attempts to ascertain the “informed” part of the “informed consent” criterion, while question R16 is meant to assess the “consent” part of the criterion. Finally, questions R17 to R19 are designed to check on the “menace of any penalty” criterion. All questions provide for answers in coded categories except for the penultimate question R20 which provides for verbatim description of the situation in the worker’s own words.

**Forced labour module [or Employment conditions module]**

**Addressed to all persons 5 years old and over working or helping in the production of palm oil of plantations**

|  |  |
| --- | --- |
| Household member No. (HM IS2) | |
| INTERVIEWER CHECK (if S13=Not manager, S15=employee or contributing family worker, S14=palm oil) | |
| 1. How many workers are in your plantation?   Categories to be defined   1. 0-5 employees 2. 6-10 employees 3. 10-50 employees 4. 51-100 employees 5. Over 100 |  |
| 1. Did you work in any other plantation during the last week? 2. Yes (go to R3) 3. No (go to R4) |  |
| 1. In how many other plantations of each type? 2. \_\_\_\_\_ <47 hectare plantations 3. \_\_\_\_\_ >=47 hectare plantations |  |
| 1. How did you find this job? (options to be decided with stakeholders) 2. Through an agent 3. Internet 4. Family member 5. … |  |
| 1. Where were you recruited? 2. In my place of origin 3. In the place of work 4. Elsewhere: \_\_\_\_\_\_\_\_\_\_ |  |
| 1. Did you pay to obtain this job?    1. Yes    2. No (go to R7) |  |
| 1. Did you incur a debt to pay for the recruitment fees?    1. Yes    2. No (go to R10) |  |
| 1. With whom?    1. recruiter/intermediary    2. employer    3. friends, relatives    4. bank    5. micro-credit group    6. other, please specify |  |
| 1. Do you have written paper on the plan for reimbursement? 2. Yes 3. No |  |
| 1. Did you sign a contract at the time of your recruitment?    1. Yes    2. No |  |
| 1. Did you sign a contract when you started the job?    1. Yes    2. No   (if R10=Yes and R11=Yes, ask R12. If not, go to R13) |  |
| 1. Were these contracts the same?    1. Yes    2. No |  |
| 1. At the time of recruitment, did you receive information about: 2. the type of job you were going to do? 3. Working hours? 4. Number of leave days per week? 5. Working conditions? 6. Wages? 7. Deductions? 8. Length of time you would be in this job? | 1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_] |
| 1. When did you start this job?   ……. / …… (MM/YY) |  |
| 1. How long do you plan to stay with this employer?   Till : …./………. (MM/YY) |  |
| 1. Are you performing any activity in your job or is there any condition of work to which you did not agree? [check as many as apply] 2. Job itself (Unfree recruitment at birth or through transaction such as bonded labour)? 3. Involuntary overtime or on-call work (with or without compensation)? 4. Involuntary work in hazardous conditions with or without protection? 5. Work in illicit activities or with use of illicit substances without consent? 6. Work at substandard or with no wages? 7. Work under degrading living conditions linked to the job? 8. Work for other employers than agreed? 9. Work for longer period of time than agreed? 10. Work with no or reduced freedom to terminate work contract?   If at least one “Yes”, go to R14 If not, go to R15 | 1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_] |
| 1. Can you refuse to do these activities? 2. Yes 3. No |  |
| 1. If you decide to stop working with this employer, can you leave when you decide (after legal notice) without negative repercussions or any risks? 2. Yes 3. No   If at least one “Yes” in R16 or R18, go to R19  If not, stop questionnaire. |  |
| 1. Why do you think you were obliged to do these activities you did not agree or that you cannot leave your employer when you want? [check as many as that apply] 2. Fear of threats and violence against me or my relatives 3. Because I am under constant surveillance 4. Because I am in an isolated place, with no access to outside world 5. Because my debt is not paid back 6. Because I do not have access to my ID card] 7. Because I do not have access to my passport 8. Because I would lose all due wages 9. Because I am an undocumented migrant and would risk deportation 10. Because I need a salary 11. Because I would have nowhere to go 12. Nothing will happen | 1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_]  1. Yes [\_\_] 2. No [\_\_] |
| 1. Can you please explain in your own words the situation you faced (VERBATIM)? |  |
| 1. When did this situation last happened?   \_\_\_\_ / \_\_\_\_\_ (MM/YY) |  |

* Child labour module

The proposed child labour module is to be implemented at the end of the questionnaire, after the employment module and the forced labour module. It consists of ten questions addressed to children 5 to 17 years old identified in the employment module to be have worked or helped last week or last twelve months in the production of palm oil. The purpose of the child labour module is to measure the prevalence of child labour in palm oil plantations and to obtain information on their working conditions.

In line with the ILO Resolution concerning statistics of child labour,[[8]](#footnote-8) child labour is defined to include all persons aged 5 to 17 years who, during the reference period, were engaged in one or more of the following categories of activities (wordings slightly modified to be consistent with the concepts and terminologies of the 19th ICLS *Resolution concerning statistics of work, employment and labour underutilization*):

* 1. *worst forms of child labour,* including forced labour and work which, by its nature or the circumstances in which it is carried out, is likely to harm the health, safety or morals of children, i.e., work which exposes children to physical, psychological or sexual abuse; work underground, under water, at dangerous heights or in confined spaces, work with dangerous machinery, equipment and tools, or which involves the manual handling or transport of heavy loads, work in an unhealthy environment which may, for example, expose children to hazardous substances, agents or processes, or to temperatures, noise levels, or vibrations damaging their health, work under particularly difficult conditions such as work for long hours or during the night or work where the child is unreasonably confined to the premises of the employer;
  2. *employment and other work within the SNA production boundary below the minimum age specified for the kind of work performed*.The minimum age for admission to employment or work should not be less than the age of completion of compulsory schooling and, in any case, not less than 15 years. Work within the SNA production boundary includes in addition to employment, unpaid trainee work, volunteer work in market and non-market units producing goods and own-use production of goods.
  3. *hazardous unpaid household services by children,* includes own-use production of services and volunteer work in household producing services for others performed under conditions corresponding to those defined under (a) above, i.e., for long hours, in an unhealthy environment, involving unsafe equipment or heavy loads, in dangerous locations, and so on. Own-use production of services refers to production of domestic and personal services by the child for consumption within the child’s own household (commonly called “household chores”).

The proposed draft module presented below has been prepared on the basis of ILO model questionnaire on the measurement of child labour developed by SIMPOC and implemented in a large number of countries throughout the world in the last twenty years or so. The proposed module contains 8 questions (C01 to C08) for assessing hazardous conditions of work by children and 2 questions (C09 and C10) for the measurement of “household chores” and the number of hours per day of the week. As in the case of the forced labour module, the question wordings, as well as the answer categories and format of the proposed child labour module are subject to review and field test by DOSM before finalization.

|  |  |  |
| --- | --- | --- |
| **Serial No in A1** | |\_\_\_|\_\_\_| |  |
| **Name of household member** |  |
| **Age of household member** | |\_\_\_|\_\_\_| |  |
| **C01. Did you have any of the following in the past 12 months because of your work?** (*Read each of the following options and mark “YES” or “NO” for all options****)***   1. Superficial injuries or open wounds 2. Fractures……….. 3. Dislocations, sprains or stains… 4. Burns, corrosions, scalds or frostbite 5. Breathing problems.................. 6. Eye problems............................ 7. Skin problems… 8. Stomach problems / diarrhea … 9. Fever……….. 10. Extreme fatigue........................................... 11. Other (specify)......................... | **Y= YES N=NO** | If all **“NO”**  **→ C04**  **Otherwise → C02** |
| 1|\_\_\_| 2|\_\_\_| 1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_| |
|  |  |
| **C02. Think about your most serious illness/injury, how did this/these affect your work/schooling?**   1. Not serious- did not stop work/schooling. 2. Stopped work or school for some time | 1|\_\_\_| 2|\_\_\_| | If 1.go to C04  If 2 go to C03 |
| **C03. For how many days did you stopped work or school due to your illness/injury?**  Specify the exact number of days --- | |\_\_\_|\_\_\_|\_\_\_| |  |

|  |  |  |
| --- | --- | --- |
| **C04. Do you carry heavy loads at work?**  1. Yes..............................................  2. No................................................ | 1|\_\_\_| 2|\_\_\_| |  |
| **C05. Do you operate any machinery/heavy equipment at work?**  1. Yes……………………….…..  2. No……………………….…… | 1|\_\_\_| 2|\_\_\_| | **→ C06**  **→ C07** |
| **C06.What type of tools, equipment or machines do you use at work?**  **(***Write down 2 mostly used***)** | 1……………  2…………… |  |
| **C07. Are you exposed to any of the following at work?**  (R*ead each of the following options and mark “YES” or “NO” for all options****)***     1. Dust, fumes, 2. Fire, gas, flames............. 3. Loud noise or vibration...... 4. Extreme cold or heat 5. Dangerous tools (knives etc) ...... 6. Work underground.................. 7. Work at heights........................ 8. Work in water/lake/pond/river…………. 9. Workplace too dark or confined ................. 10. Insufficient ventilation............... 11. Chemicals (pesticides, glues, etc.) ... 12. Explosives................................... 13. Other things, processes or conditions bad for your health or safety (specify)......................... | **Y= YES N=NO** |  |
| 1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_| |
| **Other** (specify) |  |
| **C08. Have you ever been subject to the following at work?** (R*ead each of the following options and mark “YES” or “NO” for all options****)***   1. Constantly shouted at ………………. 2. Repeatedly insulted…….. 3. Beaten /physically hurt… 4. Sexually abused, touched, done things that you did not want) 5. Other (Specify)……………… | **Y= YES N=NO** |  |
| 1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_|  1|\_\_\_| 2|\_\_\_| |
| Other (specify) |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Household Tasks** | | | | | | |
| **Serial No in A1** | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| | **Skip to Question** |
| **Name of household member** |  |  |  |  |  |  |
| **Age of household member** | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| | |\_\_\_|\_\_\_| |  |
| **C09. During the past week did you do any of the tasks indicated below for this household?** (R*ead each of the following options and mark “YES” or “NO” for all options****)***   1. Shopping for household.... 2. Repair any household equipments 3. Cooking........................................... 4. Cleaning utensils/house.................. 5. Washing clothes........................... 6. Caring for children/old/sick.......... 7. Other household tasks..... | **Y= YES N=NO** | **Y= YES N=NO** | **Y= YES N=NO** | **Y= YES N=NO** | **Y= YES N=NO** | **Y= YES N=NO** | If any “**YES**” **→C10**  Otherwise **END** |
| 1|\_\_\_|  2|\_\_\_|  3|\_\_\_|  4|\_\_\_|  5|\_\_\_|  6|\_\_\_|  7|\_\_\_| | 1|\_\_\_|  2|\_\_\_|  3|\_\_\_|  4|\_\_\_|  5|\_\_\_|  6|\_\_\_|  7|\_\_\_| | 1|\_\_\_|  2|\_\_\_|  3|\_\_\_|  4|\_\_\_|  5|\_\_\_|  6|\_\_\_|  7|\_\_\_| | 1|\_\_\_|  2|\_\_\_|  3|\_\_\_|  4|\_\_\_|  5|\_\_\_|  6|\_\_\_|  7|\_\_\_| | 1|\_\_\_|  2|\_\_\_|  3|\_\_\_|  4|\_\_\_|  5|\_\_\_|  6|\_\_\_|  7|\_\_\_| | 1|\_\_\_|  2|\_\_\_|  3|\_\_\_|  4|\_\_\_|  5|\_\_\_|  6|\_\_\_|  7|\_\_\_| |
| Other (Specify) |  |  |  |  |  |  |
| **C10. During each day of the past week how many hours did you do such household tasks?** (Record for each day separately)  1. Monday………  2. Tuesday……….  3. Wednesday……….  4. Thursday…………………….  5. Friday………………………..  6. Saturday…………………..  7. Sunday………………………. | |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_| | |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_| | |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_| | |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_| | |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_| | |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_|  |\_\_|\_\_| | **END** |
| TOTAL | |\_\_|\_\_| | |\_\_|\_\_| | |\_\_|\_\_| | |\_\_|\_\_| | |\_\_|\_\_| | |\_\_|\_\_| |

1. **Follow-up activities**

* Pilot survey

It is envisaged to conduct a pilot survey in the middle of the month of May to field test the sample selection procedure and the survey questionnaire. For this purpose, a small sample of 15 government and private estate plantations from frame I and another 15 independent smallholder plantations from frame II were selected during the mission. Table 9 shows the distribution of the pilot plantations by state and type. It should be mentioned that the pilot survey is planned to be conducted only on the basis of 9 plantations from frame I and 9 plantations from frame II. However, 6 additional plantations from each frame were selected as reserve in case it is necessary to replace a plantation or to expand the scope of the pilot test.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **9. Sample plantations for the pilot survey** | | | | |
|  | Frame I | | Frame II | Total |
| Government estate | Private  estate | Independent smallholder |
| Johor | 0 | 1 | 2 | 3 |
| Melaka | 0 | 1 | 0 | 1 |
| Pahang | 1 | 0 | 1 | 2 |
| Perak | 0 | 1 | 1 | 2 |
| Terenganu | 1 | 0 | 0 | 1 |
| Sabah | 1 | 4 | 5 | 10 |
| Sarawak | 1 | 4 | 5 | 10 |
| Selangor | 0 | 0 | 1 | 1 |
| **Total** | **4** | **11** | **15** | **30** |

One purpose of the pilot test is to obtain information on various parameters of the sample design, in particular,

* Number of plantations that could be matched with their corresponding EBs (match rate)
* Number of plantations linked to one EB
* Effectiveness of the paired EB in terms of drawing additional palm oil workers into the sample
* Number of workers with whom interview could not be conducted due to absence or refusal (non-response rate)
* Appropriate choice of the sample-take (number of sample workers to be interviewed) per linked or paired EB
* Feasibility of the adaptive cluster sampling with EBs and the appropriate choice of the number of starting points

Another major purpose of the pilot survey is to field test the survey questionnaire, in particular,

* Test the screening for identifying living quarters with workers in palm oil production
* Test the employment module of the survey questionnaire
* Test the forced labour module
* Test the child labour module
* Main survey

After the finalization of the sample design, the sample selection of the plantations for the main survey would be carried out including the identification of the linked EBs and paired EBs in the case of frame I plantations. Then the interviewers trained would be trained for locating the EBs, screening the living quarters and administrating the questionnaire. The main survey will be executed by DOSM according a time-table and protocol established by the steering committee. The ILO will participate in overseeing the implantation of the survey and will be available for clarifications and advise where needed.

DOSM in cooperation with the ILO will assess and analyze the survey results. This will involve in particular,

* Specification of the editing rules
* Specification of the counting rules
* Preparation of the tabulation plan
* Calculation of the sampling weights
* Calculation of the sampling errors of the main indicators
* Evaluation of data quality

Subject to approval by DOSM and ILO, it is proposed that a workshop be held at the end of the project to assess the effectiveness of the survey design and ideas for improving the sample design and questionnaire design in future surveys of this kind.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Annex A**

**Excel file “Sample size.xlsx”**

.

The file calculates the sample size based on a set of parameters marked in blue in column M. Based on the current values of the parameters the sample size requirement is 12028 palm oil workers per domain (row 13, col M). The parameters can be changed as obtained in the pilot survey. The sample design considers three domains:

A. Peninsular Malaysia

B. Sabah

C. Sarawak

The sample size requirements in terms of number of plantations are found in rows 14,15,16 and columns Q and R. The values for frame I on the MPOB estates are:

A.  226 plantations

B.  241 plantations

C.  242 plantations

and for frame II on Independent Smallholders are:

A.  194 plantations

B.  167 plantations

C.  165 plantations

For sample selection, these numbers should be copied as input in the Excel file of Annex B (sheet C, rows 34,35,36 and col AC) and the Excel file of Annex C (sheet mv\_penama\_lesen (LIST SH) (2) rows 19,20,21 and col. V).

**Annex B**

**Excel file “LIST OF MPOB ESTATES BY CATEGORIES 5.xls lsx”**

.

The file generates sample selections of plantations of the MPOB estates (frame I) based on a random number generated in sheet C, row 42, col. AD of the file. The sample sizes for the three domains are specified in sheet C, rows 34,35,36 and col AC. The sample output is given in the same sheet, col Y of the file. A “1” in col. Q means the plantation is in the sample, and a “0” means it is not.

**Annex C**

**Excel file “Smallholders Sample Sample Selection 20 April 2018 for farhad.xlsx”**

.

The file generates sample selections of plantations of the independent smallholders (frame II) based on a random number generated in row 1, col. P of the file. The sample sizes for the three domains are specified in sheet mv\_penama\_lesen (LIST SH) (2) rows 19,20,21 and col. V. The sample output is given in the same sheet, col Q of the file. A “1” in col. Q means the plantation is in the sample, and a “0” means it is not.

1. Malaysian Palm Oil Board, *Malaysian Oil Palm Statistics 2015 - 35th Edition*, MPOB, Ministry of Plantation Industries and Commodities, Malaysia, 2016. [↑](#footnote-ref-1)
2. US Department of Labor, <https://www.dol.gov/ilab/reports/child-labor/list-of-goods/> [↑](#footnote-ref-2)
3. Lavallée, Pierre, *Le sondage indirect ou la méthode généralisée du partage des poids*, Brussels, 2002, Section. [↑](#footnote-ref-3)
4. Thompson, Steven K., *Sampling.* Chapter 24, “Adaptive Cluster Sampling,” Wiley series in probability and mathematical statistics, John Wiley & Sons, Inc., 1992. [↑](#footnote-ref-4)
5. European Union Agency for Fundamental Rights, *Second European Union Minorities and Discrimination Survey, Technical Report*, EU-MIDIS II, Luxembourg, 2017. [↑](#footnote-ref-5)
6. ILO *Convention concerning Forced or Compulsory Labour*, 1930 (N0. 29), adopted by the 14th session of the International Labour Conference, Geneva, 28 June 1930. [↑](#footnote-ref-6)
7. ILO *Resolution concerning statistics of work, employment and labour underutilization*, adopted by the 19th International Conference of Labour Statisticians, Geneva, 2013. [↑](#footnote-ref-7)
8. ILO *Resolution concerning statistics of child labour* adopted by the 18th International Conference of Labour Statisticians, Geneva, 2008. [↑](#footnote-ref-8)