PEN Technical Guidelines

Table of content

1 Introduction 2
2 The PEN research project 2
3 The PEN research format 4
3.1 The PEN surveys 4
3.2 The PEN prototype questionnaire 6
3.3 Selection of study area 7
3.4 Sample size and sampling 8
3.5 Data entry and submission of data to Central Data Bank 8
3.6 Survey and study area narrative 8
4 Definitions 9
4.1 Forest 9
4.2 Other land categories 13
4.3 Income 17
4.4 Forest income 18
4.5 Forest User Group (FUG) 20
4.6 Village 21
4.7 Household 21
5 A detailed guide to the PEN prototype questionnaire 22
5.1 General rules and hints 22
5.2 Getting information in the village questionnaires 26
5.3 Village Survey 1 (V1) 27
5.4 Village Survey 2 (V2) 29
5.5 Annual household survey 1 (A1) 29
5.6 Annual household survey 2 (A2) 32
5.7 Quarterly household survey (Q1-Q4) 33
5.8 Attrition (drop out) and temporary absence survey (ATA) 38
6 General guidelines on fieldwork 39
6.1 Selection of study area and random sampling of households 39
6.2 Rapid rural appraisal (RRA) 41
6.3 Village census 42
6.4 Designing the household questionnaire 43
6.5 Implementing the household survey 45
6.6 Enumerators 49
6.7 Valuation and pricing 50
6.8 Security and safety 52

1 Introduction

PEN aims to collect high quality and comparable data from a variety of tropical and subtropical forest settings to study the forest-poverty interactions. At the core of the PEN project is a detailed household income accounting under different biophysical and socioeconomic contexts where there is a significant interaction between humans and forests.

PEN's success depends critically on three factors:

1) Consistency: Conducting research across different locations involving researchers with different disciplinary backgrounds and research interests presents considerable challenges. Inconstancies will arise unless deliberate efforts are made to come up with a common framework, definitions, sampling, survey design, data collection methods, pricing, coding, data entry, etc. Besides problems with definitions, inconsistencies could also arise from differences in field data collection procedures that researchers use. The purpose of the present guidelines is to contribute to consistency by outlining common research methodologies, definitions and interpretations of the PEN prototype questionnaire.

2) **Quality:** The research output and the policy conclusions derived from PEN are not credible unless they are based on high quality data and a sound analysis. While this may sound obvious, poor data are often being used to draw strong conclusions and advocate particular policy prescriptions. Ensuring high quality data means carefully designing every step of the research process. The present guidelines are also meant to contribute towards this end, although covering all relevant aspects would go beyond the scope of the guidelines.

3) **Completeness:** The Common Data Bank (CDB), which is the sum of data from the individual case studies done in the PEN format, will be a crucial resource in comparative analyses of the forest-poverty relationship. The utility of the CDB depends on it being complete, i.e., that all the cases have information on all variables included in the database. PEN partners should therefore complete all sections of the PEN prototype questionnaire.

2 The PEN research project

Note: The reader should also consult the more comprehensive project description on the PEN Web site, as well as a forthcoming working paper.

One of the least understood areas of international development is the role tropical forests play in poverty alleviation. Forests and other natural resources are crucial to the livelihoods of millions of poor people worldwide. The World Bank and FAO are arguing that forests can and must take a far greater role in meeting the UN Millennium Summit target of halving extreme poverty by 2015 (World Bank, 2001; FAO, 2005)[1]. But, how important are forests for poverty alleviation? To what extent do forests help lift people out of poverty, or are they mainly useful as gap-fillers and safety nets preventing extreme hardship? How do different management regimes and policies affect the benefits poor people derive from forests?

Oksanen *et al.* (2003)[2] conclude that forests are often ignored in Poverty Reduction Strategy Papers (PRSP) for three reasons: (i) lack of basic data on forest-poverty relations, (ii) the weak

---

[1] World Bank. 2001. *A Revised Forest Strategy for the World Bank Group. Draft 30 July 2001*. Washington D.C.: World Bank; and FAO. 2005. *State of the World's Forest 2005*. Rome: FAO.

[2] Oksanen, T. and B. Pajari and T. Tuomasjukka (eds.) 2003. *Forests in Poverty Reduction Strategies: Capturing the Potential.* EFI Proceedings No. 47, 2003. European Forest Institute, Finland.

understanding among decision-makers of the links between forestry and poverty alleviation, and (iii) lack of concrete proposals for policy reforms and investment.

The PEN project will develop general methods and subsequently generate basic empirical data on the forest-poverty relationships. PEN aims to fill the gap in knowledge through the systematic collection of socio-economic data in a variety of tropical ecosystems, using similar definitions and methodologies. This will potentially make PEN the most comprehensive global comparative and quantitative reviews of the role of tropical forests in poverty alleviation.

PEN also represents a new and innovative way of doing research. The core of PEN is the individual field studies, in most cases done by PhD students. The best field studies are often done as part of a PhD thesis. At the same time the value-added of the individual studies can be substantially enhanced by using standardized definitions and methods, which permit comparative analysis.

The project will provide well-researched advice for enhancing the way natural resource and poverty-related policies and projects are designed and implemented. If the research can document that forests and forest products are significant to rural livelihoods and household incomes, it implies that forest income must become better integrated in policy formulation such as PRSP, and also that certain forest-income related items must be included in Living Standard Measurement Surveys (LSMS - the predominant data source for PRSP). The research may also identify areas for policy reforms and investment opportunities. Alternatively, if research shows forests and forest products to be of little importance to the poor, results would help to avoid over-investment in the forest sector. Over the last two decades, the discovery of the growing importance of remittances and off-farm incomes radically changed global understanding of the nature of rural livelihoods. To bring forest income to the PRSP policy agenda could potentially be a second revolution challenging mainstream policies towards rural poverty alleviation.

The purpose of the research project is thus to *quantitatively* document forest-dependency, but due to the extent of the field work, there is also an important qualitative part. The aim is to improve forest policy formation and implementation, in order to more appropriately integrate forests and forestry in poverty alleviation to the benefit of forest-dependent people. The project's research data and results will contribute to a CIFOR managed Common Data Bank (CDB) on the forest-poverty link. The CDB will make possible global and continental level analyses of forest-poverty relationships. Local level empirical data generated by the project will thus be used to conduct analyses providing the foundation for major policy changes. By working closely with other CIFOR initiatives, the project will ensure access to a large international research network and wide dissemination of project results.

The overall research questions for PEN are: (1) **What is the current role of forests in poverty alleviation?** (2) **How can that role be enhanced through better policy formulation and implementation?**

An important aspect to investigate is the role of forests and forest products in rural livelihoods, i.e., their role in regular consumption and gap filling, insurance against shocks, and as possible stepping stones out of poverty.

PEN will also look at several potentially causal dimensions and conditioning factors of the forest-poverty link:

- the role of forests and forest products for different groups (degrees of poverty, correlation with age, sex, household headship, migration, etc.);
- the role of forests and forest products in different forest environments (forest abundance, condition, type, management, etc.);
- the role of forests and forest products in different institutional contexts (e.g., tenure regime and local management);

- the role of forest and forest products in different market contexts (market access integration, degree of competitiveness in forest product markets, etc.).

By carefully investigating how these dimensions determine the forest-poverty relationship, PEN can answer several more specific research questions:

- What is the relationship between forest use/dependency and household income/assets in different forest environments (similarities and differences)?
- What are the common characteristics behind scenarios of elevated forest income across the different environments?
- Are the poor universally more forests dependent than the better-off households?
- Does local forest ownership and management go along with higher local forest benefits (and for which groups)?
- How does forest dependency shift (e.g., from products to services) with economic development and market integration? (Although this question ideally requires longitudinal data, cross-sectional data collected in PEN can also be used.)
- Is forest dependency just a temporary phenomenon that disappears when other opportunities arise?
- How do forest product markets shape the people-forest relationship (investments, management, degradation)?
- Are there special factors (in the form of market and policy failures) about forests that prevent utilizing them more for the purpose of poverty alleviation?

3 The PEN research format

3.1 The PEN surveys

This section gives the broad overview of the PEN research format. The PEN data collection consists of three types of quantitative surveys (in addition to an attrition and temporary absence survey (ATA), see section 5.8):

1. Village surveys (V1, V2)
2. Annual household surveys (A1, A2)
3. Quarterly household surveys (Q1, Q2, Q3, Q4)

The **village surveys** (V1-V2) will collect data that are common to all or show little variation among households. It is more cost effective to collect these at the village level, and we have tried to move as many questions as possible up from the household to the village survey. There are two different village surveys, one to be done at the beginning of the fieldwork to get background information on the villages (V1), and a second one to be done at the end of the fieldwork period (V2) to get information for the 12 months period covered by the surveys.

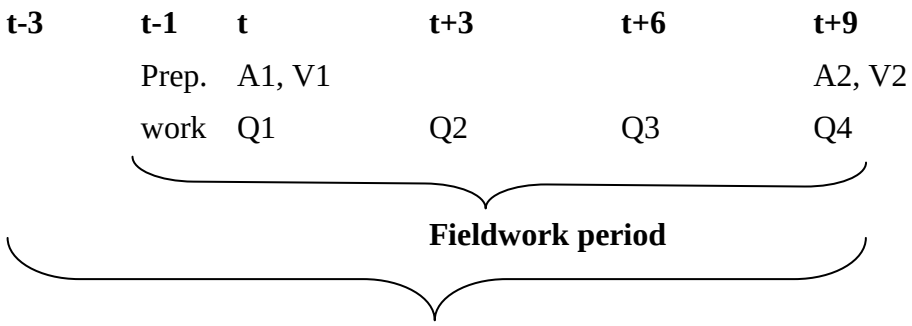
The **household surveys** are grouped into two categories. First, there are four quarterly household surveys to collect income information, while all other household information is collected in two household surveys, one at the beginning and one at the end of the fieldwork (since done almost a year apart they are labelled 'annual household surveys' to distinguish them from the quarterly household surveys). The **first annual household survey (A1)** will provide basic household information (demographics, assets, forest-related information) to be done at the beginning of the survey period (although the questions here are more time independent than in the other surveys, and some flexibility in the implementation of A1 is therefore permitted). The **second (A2**) will collect information for the 12 month period covered by the surveys (e.g., on risk), and is therefore to be done at the end of the survey period.

A critical feature of the PEN research project is to collect high quality data on forest use. This will be done through **quarterly household surveys** (Q1-Q4). After long deliberations, it was decided

that all types of income (and not just forest income) should be collected on a quarterly basis. This gives a more consistent platform for comparing various income sources, and the recall problems may be as large for some non-forest incomes as for forest incomes. The extra work in collecting information on, for example, agricultural income (often limited to 1-2 seasons) is modest.

There are two major reasons for collecting **income information on quarterly basis** rather than using a one year recall period. First, studies have demonstrated that the **accuracy and reliability** increase dramatically when the recall period is shortened, particularly for irregular income sources such as forest extraction. Thus, a prerequisite for participating in PEN is that one is willing and able to do quarterly household income surveys.[3] Second, (forest) income often has considerable **seasonal variation**. Documenting these variations can help us understand to what extent forests act as seasonal "gap fillers".

The surveys are to be designed and timed such that they cover one full year, that is, a full 12 months period (normally *not* one calendar year). **It is crucial that income and risk information in the surveys (V2, A2, Q1-Q4) refer to the same 12 months period**. This can be achieved by the following timing of the various surveys. (Time *t* refers to the start of the surveys, *t+3* refers to 3 months after the start-up, and so on.)



In addition to the quantitative surveys, one should factor in at least one month of preparatory work; without previous knowledge of the area and the language, more time may be needed. Activities at this stage will include: Rapid Rural Appraisal (RRA), possibly census, sampling, pre-testing of questionnaire, training of the enumerators, etc. (see chapter 6 of the guidelines for an elaboration of these activities).

Note 1: By doing Q1 together with A1, one will save approx. three months of fieldwork.

Note 2: All the sections of A1 *may* be asked in the 2. or 3. quarterly survey, in order to avoid too much pressure on the respondents during the first survey round. But, note that if choosing to delay the implementation of A1 it should be done for all households, as for example their savings and debt position may show seasonal variation.

Given this sequence, it may be possible to do a complete PEN fieldwork within a time span of approx. 10 months. However, unforeseen logistical problems could easily prolong this period – *Murphy's Law is particularly relentless in the tropics!* The researcher may not have to be in the field for the full period, but the need for supervision of enumerators, checking of questionnaire, further RRA and key informants interviews, qualitative surveys, etc. suggest that he or she most of this period should be in the field.

---

[3] See PENNEWS 2/2005, which reports on an experimental study in Nepal, showing that the forest use reported in monthly (and quarterly) surveys were almost twice the figures reported in annual recall surveys.

All PEN partners are recommended to make their own more detailed time table of activities.

3.2 The PEN prototype questionnaire

The PEN surveys referred to above are those given by the PEN prototype questionnaire, plus any additional questions and sections included by the PEN partner for his or her particular research topic.

The PEN prototype questionnaire asks for all the information to be included in the Common Data Bank (CDB) and the global analysis. Thus PEN partners are expected to use this as a mandatory component of their surveys. Indeed, PEN partners are defined as those who collect a data set for the CDB. *It is essential for the success of PEN to get a data set as complete as possible from each study site.*

In addition, most researchers will extend the PEN prototype questionnaire to fit his or her particular research topic. The questionnaires to be applied in the field will therefore include questions and sections that reflect the particular focus of the PEN partner. For example, if the focus is on forest product markets, it is natural to include new sections that go into the details of marketing. This additional information will *not* be included in the CDB.

As explained in the introduction, a key for creating a reliable CDB is that data are collected in the same format. The following rules therefore apply to any modification of the PEN prototype questionnaire:

1. It is essential to keep the exact meaning of the question. While a word for word translation rarely will work, but both the PEN partner and enumerators should work through these guidelines to get the intention behind the questions (see section 5 of the guidelines) and also pretest the questionnaire to see how to best ask the questions in the local language.
2. Rather than to modify the standard questions, PEN partners should create *supplementary* questions (and sections) to capture site specific information and their special research needs.
3. Exceptions to this rule, i.e., the alteration or omission of PEN questions due to the local irrelevance or inadequacy of PEN standard questions, invariably need to be pre-authorised by the PEN coordinator.
4. In some cases one may ask more detailed (disaggregated) questions than the PEN prototype questionnaire does, e.g., on forest categories. But, one should make sure that the information collected can be aggregated into the categories used in the PEN prototype questionnaire.
5. In the sections about forest and agricultural products, one should add the locally most common products, and also leave space for additional ones (see more detailed suggestion in section 5).

Having said this, it is also necessary to stress that the questionnaire cannot just be printed out from the PEN Web page and applied straight away. In fact, some parts of the questionnaire – the quarterly survey in particular – is more of a data worksheet than a survey instrument to be read out to the respondents. The questionnaire is in its language designed to be understood as precisely as possible by the PEN student and the enumerators. It is not designed to be directly read out to (and understood) by households being interviewed; some "translation" to their common terminology has to be done.

The importance of **pretesting the questionnaire** will be stressed several places. This is needed to become acquainted with the questionnaire, find out how to best ask the questions (e.g., local terms), get an overview of common products in the study area, and identify and formulate additional questions for your particular needs.

3.3 Selection of study area

The selection of study area is a critical decision. The area is often chosen due to certain features of interest for the research project: a minimum level of forest dependence (a *must* for all PEN studies), and a few characteristics relevant for the individual study (going beyond the core of PEN).

In addition to these, there are two concerns that should be underlying the decision about the selection of both study area and villages within that area: **representative and variation**.

It is important to strive for representative areas, and try to avoid special cases, e.g., areas with unusually valuable forest products, unusually favourable or unfavourable conditions for income generation, or a history of very heavy donor intervention. After all, the individual PEN studies would like to do analysis and derive conclusions that have a bearing beyond the study area and villages. Similarly, the PEN project intends to arrive at some generalized conclusions about the forest-poverty relationship in developing countries.

At the same time, we also realize that PhDs and other research projects often are driven by an interest to investigate a particular phenomenon: conflicts around a national park, a new scheme for biodiversity payment, a pilot project of local forest management, the exploitation of a particularly valuable forest product, etc. But, in order to investigate the impact of programmes, policies or special conditions, one must have variations in these variables (control groups).

Cavendish (2003)[4] discusses different criteria that can be used in choice of study area and villages within that area to ensure such variation in data: forest type, forest tenure, local agent, source of risk, levels of poverty, market distance, population density, and migrants. While it may be difficult to find variation along all these dimensions within the study area, one should think carefully through the selection criteria for the study area and villages within that area. For example, assume that you would like to study how market access and forest resource availability affects the use of and dependence on forests. You should then choose a study area where you find variation in both market access and forest resources. You may create a two-way table (four cell matrix) with good-poor market access, and high-low forest resources. The village sampling should then try to identify villages within each of these four cells. (Admittedly, this may be easier said than done as good market access and low forest resources often are positively correlated.)

More generally, in PEN research we would like to *get variation* within the study area along the following dimensions:

- Market access/remoteness, which has an important bearing on relative prices and off-farm employment opportunities, and population densities.
- Forest abundance, and possibly also types of forest.
- Income (poverty) levels, although there will be some natural variation within villages.
- Institutional aspects, e.g., management regime, and effectiveness of local forest management.

While it may be difficult to get variation in all these, market access and remoteness is often a critical variable. It is therefore key criterion to be used by most studies for selecting study area and sample villages (i.e., choosing villages along a gradient of varying market access and remoteness).

Section 6.1 discusses general sampling issues further.

[4] Will Cavendish. 2003. *How Do Forests Support, Insure And Improve The Livelihoods Of The Rural Poor? A Research Note.* Unpublished paper, CIFOR. Available at the PEN Web-page.

3.4 Sample size and sampling

To be included as the PEN data set, a minimum sample of 100 households is required. Since there will always be some default (e.g., drop-outs during the household survey), a "leakage" of at least 10 % has to be factored in, so the minimum target sample has to be 110 or possibly higher. Having said this, we consider this to be an absolute minimum, and in most cases the sample should be well above this. 200-300 households would often be a reasonable sample, but this obviously depends on the time and resources available to the individual PEN partner and the type of analysis to be done. Yet, it is in the self-interest of the individual studies to have a reasonably large sample size to permit intermediate and detailed levels of statistical analysis, e.g., compare sub-samples. A larger sample will also ensure more variation and therefore make it possible to test more hypotheses. These considerations should be just as important as the mandatory minimum sample required by PEN.

The general rule is that households should be selected through random sampling from the total population in the village. In other words, no *ex ante* sampling criteria should be imposed, such as respondents being farmers or having forest income. Indeed, a key question is: "who are the forest dependent?", and information about non-forest dependent household are also needed to address that question.

See chapter 6.1 of the guidelines for a further discussion of selection of study area, villages and households.

3.5 Data entry and submission of data to Central Data Bank

A data entry template in MS Access is available (together with the user manual) on the Pen web site. The Access screen view visually resembles a questionnaire, which makes it easier to fill in information from the questionnaires. The data is stored in a table (spreadsheet-like file) in the background, which can be transferred to any statistical programme for analysis.

Using a standard template for data entry has several advantages. First, it will be much faster compared to entering data directly into a spreadsheet. Second, the data will be in a uniform format, easy to read by a statistical programme and to use in later data analysis. Third, it will minimize the data entry errors, both by using a questionnaire-like template and by including logical checks.

After completion of all the surveys and data entering, the data set is submitted by the PEN partner to CIFOR for checking and approval.

3.6 Survey and study area narrative

In addition to a PEN data set, the submission should also include a short description (narrative) of the survey and the study area. The narrative will provide important background information for those using the data in the global analysis. This introductory information can be used directly by the PEN student in forthcoming PhD theses and papers, as well as helping CIFOR and others in interpreting the results of the statistical analyses. The length and format can therefore to some extent be adapted to the particular need of the individual PEN studies, in order not to impose too much additional work.

The suggested format of the narrative is 10-15 pages, depending on sample size and the variation in the study area, among other things. It should provide information complementary to what is included in the PEN prototype questionnaire. Here is the suggested format of the narrative:

PEN NARRATIVE

1. Purpose of project (objectives/research questions, summary) - 0.5 page
2. The survey
    a. Format: sampling (criteria used), number of households, timeline (start-end of each survey), any additions to PEN prototype questionnaire, etc.
    b. Implementation, problems encountered, enumerators, etc.
    c. Explanations of the local units used, and their respective conversion factors.
3. Study area characteristics (including all villages in the survey)
    a. Brief history of the region
    b. Demographics: ethnic composition, migration patterns, etc.
    c. Major economic activities (e.g., agricultural systems, timber trade, remittances)
    d. Seasonal calendar: major activities throughout the year (by month)
    e. Markets and market access
    f. Forest product markets
    g. Major land cover and land uses: biophysical description and recent changes
    h. Description of conservation areas: size, protection status, permitted uses, degree of enforcement.
    i. Tenure institutions: a more qualitative description of institutions governing the use of land, forests and natural resources
    j. Government and other development/conservation projects
    k. Any calamities (e.g., drought, fire, economic crisis, war, famine) in the last X years that must be known to fully understand the data gathered?
    l. Other relevant issues
4. The villages
    - Use the same headings as under 2, but provide village-specific, more disaggregated information.

Appendix: Copy of the actual questionnaire applied in the field.

4 Definitions

This section presents the key definitions for use in PEN research, particularly when using the PEN questionnaire to collect data. It should be noted that the definitions presented here are for the purpose of ensuring consistency among all the PEN cases studies. Everyone should carefully adhere to the same definitions so that the data of the various case studies can ultimately be aggregated, compared, and contrasted. There are alternative definitions that could legitimately be applied, but these are not to be used in the PEN case studies.

4.1 Forest

4.1.1 Introductory considerations

The most commonly used definition of *forest* is the one provided by FAO[5]: "Forests are lands of more than 0.5 hectares, with a tree canopy cover of more than 10 %, which are not primarily under agricultural or urban land use… The trees should be able to reach a minimum height of 5 meters in situ."[6]

A further distinction is made between natural forest and forest plantation. A natural forest is defined as "forests composed of indigenous trees, not planted by humans". *Plantation forest*, on the other hand, is defined as: "Forest stands established by planting or/and seeding in the process of afforestation or reforestation. They are either: of introduced species (all planted stands), or intensively managed stands of indigenous species, which meet all the following criteria: one or two species at plantation, even age class, regular spacing."[7]

In addition to the establishment (natural *vs.* planted), forests can be further classified along a number of other dimensions, including:

1. The degree of **human disturbance** and modification (management and intensification, see below);
2. The canopy closure (cover). FAO makes an important distinction between open and closed forests: **closed forests** have a canopy cover above 40 %, whereas **open forests** have a canopy cover between 10 and 40 %.
3. **Type of trees**, for example, between broad leaved (hardwood), coniferous (softwood), bamboo/palm and mixed forests, or between native, exotic and mixed forests (relevant for plantations).
4. **Climate (rainfall),** where a distinction between **wet (humid)** and **dry** forest is common, and sometimes with semi-wet as an intermediate third category.
5. **Management objective**, where a broad distinction is made between protection and production forest, but with the possibility of several sub-categories reflecting a continuum in the permitted uses and degrees of management.[8]
6. **Ownership (tenure)**, with state, private and communal forests being major categories.

Classifications along these dimensions are useful for different purposes, and most of them (except no. 3) are discussed below and in some way included in the PEN research format.

---

[5] See: FAO (2000). *FRA 2000. On definitions of forest and forest change.* Forest Resources Assessment WP 33, Rome, 2 November 2000.
http://www.fao.org/documents/show_cdr.asp?url_file=/docrep/006/ad665e/ad665e00.htm

[6] The full definition reads: "Forest includes natural forests and forest plantations. It is used to refer to land with a tree canopy cover of more than 10 percent and area of more than 0.5 ha. Forests are determined both by the presence of trees and the absence of other predominant land uses. The trees should be able to reach a minimum height of 5 m. Young stands that have not yet but are expected to reach a crown density of 10 percent and tree height of 5 m are included under forest, as are temporarily unstocked areas. The term includes forests used for purposes of production, protection, multiple-use or conservation (i.e. forest in national parks, nature reserves and other protected areas), as well as forest stands on agricultural lands (e.g. windbreaks and shelterbelts of trees with a width of more than 20 m), and rubberwood plantations and cork oak stands. The term specifically excludes stands of trees established primarily for agricultural production, for example fruit tree plantations. It also excludes trees planted in agroforestry systems." See:
http://www.fao.org/documents/show_cdr.asp?url_file=/DOCREP/004/Y1997E/y1997e1m.htm

[7] http://www.fao.org/documents/show_cdr.asp?url_file=/DOCREP/004/Y1997E/y1997e1m.htm

[8] See FAO definitions for a finer classification.

4.1.2 PEN definitions of forest and forest categories

Forest

In general, PEN adopts the FAO definition of forest. The exclusion of urban land use is irrelevant for PEN as the surveys are done in rural areas.[9] PEN will thus apply the following definition:

> *Forests are lands of more than 0.5 hectares, with a tree canopy cover of more than 10 percent, where the trees should be able to reach a minimum height of 5 meters in situ, and which are not primarily under agricultural land use.*

Note that this definition includes both primary and secondary forest, native and exotic forests, as well as closed and open forest (e.g., woodlands). Thus PEN research should *not* use the term "forest and woodland" as "forest" already includes "woodland".

Categories of forests

We also follow the FAO distinction between natural and plantation forest. However, the FAO's "natural forest" is further subdivided according to the degree of management and human disturbance. PEN will instead use three forest categories, defined in the following way:

1. ***Natural forest*** *consists of indigenous (native) tree species. It is managed only to a very limited degree, i.e., one may practice "tolerant forest management in which the native vegetation is largely conserved or reconstructed through successional processes".*[10]

In natural forests, most beneficial trees occur spontaneously, although there may be some degree of management to stimulate the frequency and growth of these trees, e.g., by clearing competing vegetation.

2. ***Managed forest*** *consists predominantly of indigenous vegetation, and with active management to increase the frequency and productivity of beneficial species. The management will include felling (trimming, thinning in addition to regular harvesting) and planting of indigenous and/or exotic species.*

Managed forest will include both what is termed *production forest*, i.e., forests managed for timber production, and forests managed for various NTFPs. Forests and old forest fallows that have been enriched, e.g., by the planting of fruit trees, will therefore fall under this category. (See also discussion of *fallow* in the section below.)

3. ***Plantation*** *consists of forest stands established by planting and/or seeding in the process of afforestation or reforestation. They are composed either of (a) introduced species (all planted stands), or (b) intensively managed stands of indigenous species, which meet all the following criteria: one or two tree species planted, even age class, regular spacing.* (FAO definition)

Take special note of the additional requirements above which restrict the definition of plantations with indigenous (native) species. If the land area is planted primarily with native species, but the trees are of uneven age and spacing, then this forest would not be categorized as "plantation", but instead as a "managed forest".

---

[9] The more extensive definition referred to in the footnote above also includes potential reforestation areas. These are not included in the PEN forest definition.

[10] K.F. Wiersum. 1997. "Indigenous exploitation and management of tropical forest resources: an evolutionary continuum in forest-people interactions". *Agriculture, Ecosystems and Environment*, 63: 1-16.

While these three categories exist along a continuum, some important distinctions are made. Moving from category 1. to 2. is marked by the active felling and planting in order to increase the production of beneficial trees. Moving from 2. to 3. is marked by the vegetation going from being predominantly spontaneous indigenous species to being predominantly planted species (either indigenous or exotic).

In other words, our primary criterion for the distinction between forest types is the *degree of human intervention*. Our secondary criterion is *the type of tree species* that the vegetation systems are composed of.

Including plantations into the forest definition is by some seen as controversial: after all, it would almost only be the size of the cultivated plants that distinguishes these systems from agriculture. However, by clearly distinguishing in PEN surveys between the two types of natural forest and plantations, one can do separate analysis with a more narrow forest definition (excluding plantations) and a broader one (including plantations).

Land tenure

Ownership or tenure of the forest and agricultural land is an important aspect for many key research and policy questions. The land tenure regime consists of many dimensions, and the three digit coding system (*code-tenure*) is based on three aspects: (1) the formal or ***de jure*** **owners** of the forest (state, community, individuals); (2) the ***de facto*** **users** of the forest (state, community, individuals, or combinations); (3) the degree of rules **enforcement** (high, medium/low, none). This is elaborated further in the code list.

Open-closed forest

In the codes for forest categories (*code-forest*), a distinction is made between open and closed forests. This follows the FAO definitions with a 40 % canopy cover dividing line. Thus,

- **Closed forests** have a canopy cover above 40 %. Examples include tropical rainforest and mangrove forest.
- **Open forests** have a canopy cover between 10 and 40 %. Open forests generally have a continuous grass layer. Examples include the wooded savannahs and woodlands in Africa, and part of the *cerrado* and *chaco* in Latin America.

Wet-dry forests

We are not asking PEN collaborators to specify whether the forests in their case study areas are humid or dry. For that reason, we do not provide a definition here. We will be able to make this distinction at a later date on the basis of the rainfall data that will be collected in the village surveys.

Protection status

The "protection" status of the forest will not be specified in the prototype questionnaire. It should, however, be specified in the village narratives that will accompany the submission of the PEN data set.

Summary

The above discussion of forest definitions is summarized as follows:

1. PEN adopts the FAO forest definition: Min. 0.5 ha, min. 10 % canopy cover, and min. (potential) height of 5 meters, and not agriculture as the primary use.

2. PEN uses the following three categories of forest: (1) Natural forest; (2) Managed forest; (3) Plantation forest.
3. Agroforestry (or silvipasture) is not considered forest but treated as a separate land category.

4.2 Other land categories

Overview of land categories used in the PEN questionnaire

In the village questionnaire (V1, D1) and the codes for land categories (*code-land*), a number of other land categories are used in addition to the three forest categories. These are defined in this section. The main land categories used in PEN research is shown in the table below.

Category	**Brief definitions & comments**
Forests:	
Natural forest	Indigenous species with only limited management.
Managed forests	Predominantly indigenous species, and management including felling and planting of indigenous and/or exotic species.
Plantations	Forest stands established by planting and/or seeding.
Agricultural land:	
Cropland	Land cultivated with crops. But, there is a fine distinction between cropland and fallow, agroforestry or plantation, see definition in text.
Pasture	Land used for herbaceous forage crops, established by humans and/or with active management.
Agroforestry	Land use established by humans combing trees and crops, either on a spatial or temporal scale.
Silvipasture	Land use established by humans combining trees and pasture.
Fallow	Agricultural land temporarily (up to 15 years) not being used for crops or pasture.
Other land categories:	
Shrubs	Woody perennial vegetation less than 5 m in height.
Grassland	Land with grass as the predominant natural vegetation; may have scattered trees (savannah) of less than 10% canopy cover.
Residential areas, infrastructure	Land used for buildings, roads, etc.
Wetland	Land where water saturates the soil, not classified as by forests or agricultural land.
Others	Land not fitting into any of the above categories.

Note again that a major distinction in the tables and codes is made between forests and agricultural land. PEN-defined forests are landscapes with trees, where agriculture might be occasionally practiced, but is not being the predominant land use. Landscapes with trees where agriculture is the dominant land use (e.g., agroforestry or silvipasture, but not limited to these categories) are *not* forests by our definition.

Cropland

Cropland is defined as land from which crops are harvested. But, a few exceptions to this general definition are worth pointing out:

- Idle cropland, i.e., land which is normally used for crops but temporarily not cropped, should be included under *fallow*.
- Following the FAO definition, land with trees grown for wood or timber should be categorized as *plantation*.

- Annual crops cultivated in combination with trees should be categorized as *agroforest.*

Pasture

Pasture is generally referred to as land covered with grass or other vegetation (herbaceous forage crops) eaten by grazing domesticated animals. In PEN, we limit the definition of pasture to cover only land where grasses and/or legumes have been established by humans and/or involve some other form of active management. Thus, natural grassland used for grazing is *not* included in our definition of pasture, but is instead classified as grassland (see below).

Agroforestry

Agroforestry is not classified as forests (it "fails" on the agricultural use criterion), and is in PEN research a separate land category. Agroforestry is defined as "an agricultural land use that combines growing trees (woody perennials) with annual (*herbaceous*) crops, either on a spatial or temporal scale."[11] Note that agroforestry is characterized both by the type of vegetation (mix of trees and annual crops) and the intentional establishment of this system by humans for agricultural purposes.

Silvipasture

While agroforestry is the combination of trees and crops, silvipasture is the combination of trees and pasture. To qualify as a silvipasture system, either the trees or the pasture (or both) must have been established by humans. Thus, a savannah area with naturally existing trees and grass cover, used for grazing livestock, does *not* qualify for being a silvipasture (nor a pasture). As for agroforestry, land devoted to silvipasture is not considered forest in the PEN research.

	Tree canopy cover		
	None	**0-10 %**	**Above 10 %**
Planted grass (or legumes)	**Pasture**	**Silvipasture**	
Natural grass cover	**Grassland**	Savannah, but included in "**Grassland**" in PEN research	**Forest** or **silvipasture** if trees established by humans

Fallow

Fallow land refers to land which is part of an agricultural (cropping or pastoral) rotation system, but which is temporarily not being cultivated. This category raises some challenges, as fallow land in terms of vegetation cover can represent a continuum from bare land, grassland, bush/shrubs, to young forest and mature forest.

---

[11] The broader definition of the World Agroforestry Center (ICRAF) is: "Agroforestry is a dynamic, ecologically based, natural resource management system that, through integration of trees on farms and in the agricultural landscape, diversifies and sustains production for increased economic, environmental and social benefits." http://www.worldagroforestry.org/Agroforestry.asp

In PEN research the term should be used for land which is considered agricultural land, that is, it has been used for growing agricultural crops and/or pasture within the last 15 years, and is it likely to be used for cropping/livestock again in the short/medium term.

Should *fallow forest* be classified here, as a non-forest land use, or should it be classified as a forest? If it is part of a regular rotational cycle, it should be categorized as *fallow*, even if it has the biophysical forest characteristics. If, on the other hand, it was cleared a long time ago and there are no immediate plans to bring it back into cultivation, it should be classified as *forest*. "Long time ago" is in this context defined as more than 15 years, which is in the range where many consider the fallow to have developed into 'proper' forest.

In short, the rules are as follows:

- If the fallow age (= time since last crop) is above 15 years it should be classified as forest irrespective of any plans for agricultural conversion in the future (of course, provided that it satisfies the forest definition of a 5 meters tree height and minimum 10 % canopy cover; if not it should be classified as shrub, grassland or other land).
- If the fallow age is below 15 years, and there are no immediate plans for bringing it back into agricultural use, the land should be classified as forest, shrub, grassland or other land.
- If the fallow age is below 15 years, with plans for bringing it back to agricultural use (crops or pasture) within the next 5 years, it should be categorized as *fallow*.

Thus, to qualify as fallow the age should be below 15 years *and* there should concrete plans for bringing the land back into agricultural use.

Note also that if fallow is used for grazing, it should be classified as "grassland", provided it has less than 10 % canopy cover.

Shrubs

According to FAO, *shrubs* "refer to vegetation types where the dominant woody elements are… woody perennial plants, generally of more than 0.5 m and less than 5 m in height on maturity and without a definite crown."[12] Thus, woody vegetation is broadly classified into forest and shrubs, where a plant height of 5 meters draws the line. Shrubs with a closed canopy are referred to as *thicket.*

The term *bush* is popularly used interchangeably with *shrubs*, but *bush* sometimes also refer to open woodlands and wilderness areas. If, however, the canopy cover of the woodland is more than 10 % and the tree height above 5 meters, it should be grouped as *forest*.

Grassland

Grassland has naturally occurring grass as the predominant vegetation. If it has trees scattered around (and canopy cover below 10 %), it is referred to as *savannah* or *wooded grassland*, but is still categorized as *grassland* in PEN research. If the canopy cover is above 10%, it should be classified as *forests* (e.g., woodlands in Africa, which is a form of open forest with canopy cover between 10 and 40 %).

Grassland may be used for grazing domestic animals, and may then be referred to as *rangeland*. In PEN research it would still be categorized as grassland, and not pasture.

Wetlands

[12] http://www.fao.org/documents/show_cdr.asp?url_file=/DOCREP/004/Y1997E/y1997e1m.htm

Wetlands refer to land areas where water saturates the soil, either permanently (*swamp*) or parts of the year. *Mangrove* forest is often considered a wetland, but should in PEN research be included in the "unmanaged natural forest" category. If a wetland is used for agriculture, e.g. for paddy rice, it should be included in one of the agricultural land categories.

4.3 Income

Income is generally defined as the value added of labour and capital (including land). The income of a household is therefore the return to the labour and capital it owns, used in own production and income-generating activities (self-employment or business) or sold in a market (e.g., wage labour). We also include transfers in the income definition, e.g., in the form of remittances or pensions. Thus, income consists of three broad components:

1. Income (cash and subsistence) from self employment and business (sections B, C, D, E, H, I in quarterly household survey).
2. Wage income (section F) and income from renting out capital (including land) (section J, question 6)
3. Transfers (section J).

The basic income equation for income from self-employment or business (agriculture, forestry, and any other business) is:

$$I = \sum_{i=i}^{n} p_i y_i - \sum_{j=1}^{m} q_j v_j$$

Income (I) is gross value (price times quantities of all *n* products) minus the total costs (price times quantities) of all *m* purchased inputs (e.g., fertilizers, seeds, tools, hired labour).[13] Note that the costs of family labour should *not* be deducted to obtain household income. One may want to measure the amounts and costs of family labour for other purposes, e.g., comparisons of the profitability of different activities, but it is not needed for calculating household income.

Total household income is the sum of cash income and subsistence income, the latter referring to the value of products being consumed directly by the household or given away to friends and relatives. One should be aware that many respondents may consider income to mean *cash income only*. It is very important that our expanded definition of income (subsistence + cash) be clear to the enumerators and respondents.[14] We return to the issue of pricing and valuation of subsistence consumption in chapter 6.

Barter trade and gifts

Barter sale is a form of trade where a commodity (e.g., agricultural produce) or service (e.g., labour) is exchanged for another commodity or service, without any monetary transactions involved. For example, a household may exchange with the neighbour 10 bags of rice for one calf. Barter trade should be grouped together with regular sale in the PEN questionnaire, and is therefore also part of the income definition.

Gift is the transfer of a commodity or service without any direct compensation. In the PEN survey, gifts given away should be grouped along with household consumption.

---

[13] Maintenance of capital stock (or depreciations) should also be included, but this will have limited applicability for most households.

[14] The income definition used is PEN is an expanded one only compared to some popular use. In economics our definition is the standard one.

Thus, we get the following equation:

> *Total production = own use (i.e., household consumption + gifts) + sale (i.e., cash sale + barter)*

Note that smaller gifts received are *not* to be recorded in the questionnaire and counted as income for the recipient household, as it is already included in the production (and thereby income) definition of the household giving the gift. Larger gift should, however, be recorded (see section 5, quarterly questionnaire, section J).

Where do we draw the line between gifts and barter? In all cultures there are norms of *reciprocity*, that is, a kind of informal trade of commodities and labour where the trade over time should balance. "I help you with my labour or some food today, and you help me later when I'm in need of it." We still group such transfers as gifts rather than barter. To qualify as barter there must be a non-monetary payment *directly* related to the commodity or service exchanged.

4.4 Forest income

4.4.1 What are forest products and income?

Forest products can generally be defined as products collected from a forest. Forest products include timber and a wide range on non-timber products (NTFPs), both tree-based (e.g., fruits), various plants (e.g. tubers), and fauna (e.g., caterpillars or bush pig). While this definition might sound trivial, there are a few difficult cases that need some clarifications. Two key issues are: Should products that happen to be found in the forest, but that do not depend on the existence of forest, be classified as forest products? Should conventional forest products originating from outside forests (e.g., in agroforestry systems) be considered forest products?

A general principle related to the first question could be as follows: A product harvested or collected from the forest is defined as a forest product if its supply depends on the existence of the forest, i.e., it will disappear if the forest is gone. Some particular categories are worth highlighting:

1. **Minerals** from forests, e.g., gold, do not depend on the existence of forest, at least not in the relevant time perspective. In PEN research they are *not* classified as forest income, but are included in the category of non-forest environmental income.
2. **Fish** and other aquatic animals (e.g., shrimps) are another potentially important source of income, which can be caught in water bodies inside forests and/or in systems with a tight ecological relationship between the two. In some cases, such as large parts of the Amazon where fish and forests clearly are ecologically connected, a strong argument can be put forward for defining fish as a forest product. In other cases it would be more questionable. In the PEN survey we ask about fish in a separate table, and where it has been caught (land type surrounding the river/lake), in order to classify it in the later analysis. We also distinguish between fish caught from the wild (environmental income and possibly also forest income) and from aquaculture (not environmental nor forest income).
3. **Game meat** is yet another difficult category. Some animals may live most of the time in the forest, but sometimes come out to feed in agricultural fields, where they also normally are hunted. In this case they should be classified as "forest product". In other cases, there is limited connection to the forest, i.e., they live and feed mainly outside forests, e.g. in grasslands. They should then be categorized as non-forest environmental income. The questionnaire asks explicitly about where the product derives from in order to make this distinction.
4. **"Forest products" outside forests** are products which are conventionally considered forest products, but then are domesticated and grown outside forests. They can be split into two categories. The FAO forest definition makes a distinction between wooden products and agricultural products, in the way that "trees established primarily for agricultural production"

are excluded in the forest (plantation) definition. Thus, income from a timber plantation is included as forest income, while a fruit tree plantation is not, even though the fruit tree may be native to the forests of the region. This is debateable, and PEN does not need to adopt one particular definition as the product and land codes (where the product originates) used will permit analysis using different definitions.
Note: If the same product originates from different sources (i.e., different land codes), one may use several rows for the same product in the questionnaire.

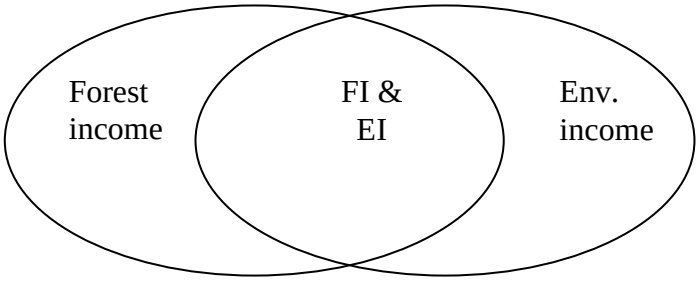
5. **Forest services:** With the increasing potential for direct local income from environmental services, it becomes important to distinguish between environmental services in general and forest services in particular. Some service payments are clearly linked to the maintenance or establishment of forests, e.g., carbon payments or payments for hydrological services. In other cases such as ecotourism forest based services may just a part of the total package. In such cases one may try to establish how much of the ecotourism income is related to forests, and this should be done in the narrative accompanying the PEN data set.

4.4.2 Forest income vs. environmental income

The ambiguity as to what is or is not a forest product (and income) points to the need for clarifying another concept, namely environmental income. PEN uses the following definition:

> *Environmental income are incomes (cash or in kind) obtained from the harvesting of resources provided through natural processes not requiring intensive management.*

From the above it should also be clear that not all forest income (FI) qualifies as environmental income (EI), and vice versa. This is illustrated in the diagram below, where we have three categories (moving from left to right in the box): (1) FI but not EI, e.g., timber from forest plantations. (2) Both FI and EI, e.g., fruits and timber harvested from (managed and unmanaged) natural forests. (3) EI but not FI, e.g., gold mining and fishing outside forests.



In the PEN survey it is important to include non-forest environmental income, partly to see the relative importance of forests to other ecosystems and natural resources, and also to get a complete estimate of total income. This is done in section E in the quarterly household survey.

4.4.3 Types of forest income

Given a definition of forest products (and services), the next step is to clarify the various types of forest income, and how to measure them. The more general issues related to "Pricing and valuation" and definition of income from self-employment are discussed in section 6.

PEN distinguishes between four types of forest income:

1. Income from self employment in the harvesting of forest **products in the raw**, and used or sold in an unprocessed way. Section B of quarterly household survey (Q1-Q4).
2. Income from self employment of **processed (value added**) forest products, e.g., woodcrafts and carpentry. Section C (Q1-Q4).

3. **Wage** income from employment in forest based activities, such as logging or tourism. Section F (Q1-Q4).
4. **Direct payment** (transfers) to the household for forest-based environmental services, e.g., carbon credits or profits from community-based forest ecotourism. Section C (A2).

The most difficult one is the second category. The basic problem is this: Should calculations of forest income be based on: (1) the value of the processed product (e.g., wood carving), or (2) the value of the forest raw material (wood) which the product is based on? What about a table the carpenter made, or the canoe made from a large log? Or the meal sold in a restaurant based on game meat from a forest?

The key issue is what we define as the forest product: the wood carving or the wood for making the carving? Given our choice, income should be calculated the normal way as discussed above, i.e., we should take gross value and deduct the costs of purchased inputs, but not household labour.[15]

If someone is involved in wood carving, carpentry, basket making, etc., both the value of the finished products should be recorded, as well as the price and value of the raw forest material used in the production (see section C of Q1-Q4). Thus we will get both an estimate of the forest product in raw, as well as the value added, and can distinguish between different types of forest income in the later analysis.

Another definitional issue is where exactly to draw the line between **unprocessed (in the raw) and processed forest products.** In PEN "processed" refers to significant modifications or changes of the product, e.g., turning wood into charcoal or a chair, or turning clay into a pot. Minor modifications, for example, cutting rattan canes or bark into smaller pieces, or washing and/or drying the product would not qualify, and the products should still be classified as unprocessed. But, arguably, products exist along a continuum, and if in doubt you should consult the PEN advisor/coordinator to ensure consistency across the studies.

4.5 Forest User Group (FUG)

A forest user group (FUG) can be defined as a group of people who use and maintain a forest, and who share the same rights and duties[16] to products and services from the forest.[17] The definition presupposes some forest with collective property rights, i.e., a FUG cannot exist – according to our definition – if all forests are privately owned. The group may or may not be formally organized, but a minimum level of organization is required, including regular meetings. It may have originated through customary law, or may have been established through outside interventions, e.g. by NGO or government. Examples of FUGs are groups specifically designed for community forest management, producer organizations that include forest products in their portfolio, grazing associations, and natural resources committees within a formal village structure.

A wider definition of a FUG includes all the users of the particular forest. It is synonymous to the group of all stakeholders involved. This interpretation is too wide for PEN purposes. A minimum level of organization and involvement in the management of the forest is needed to qualify as a

---

[15] For a further discussion, see: Sjaastad, Espen, Arild Angelsen, Pål Vedeld, and Jan Bojö. 2005. What is environmental income? *Ecological Economics* 55 (1): 37-46.

[16] Sometimes, e.g., in some community forestry groups in Southeast Asia, the managers of the FUG get a higher income from the project as a form of compensation for their work. So rights and duties may be unequal in these cases, but it would still be defined as a FUG.

[17] This definition draws on the one used by the International Forestry Resources and Institutions (IFRI) research programme of Indiana University, cf. their *Form U – User Group Forms: Guidelines*.

FUG in the PEN surveys. Not in all PEN cases will there be FUGs in place; in particular in the Neotropics they are less common than in Africa and Asia.

Being member of a FUG entails a set of rights and obligations. The rights are in the form of preferential access and use rights to the forest resources, and possibly also other benefits. The obligations are in the form of obliging to the rules and contributing to the management. The management of the forest involves two sets of tasks:

1. The establishment and/or modifications of *rules* with respect to:
    a. *Access* to the forest products, e.g., who can collect what, when, and how much?
    b. Permissible uses of forest products, e.g., only firewood extraction for domestic consumption is allowed while commercial sales are prohibited.
    c. Maintaining or enhancing the productive capacity of the forest, e.g., establishment of firebreaks.
2. Implementing and enforcing the rules, e.g., policing, punishment, and organizing joint work.

A FUG may also be involved in a number of other activities, e.g., working together on harvesting, processing and marketing of forest products.

In summary, for PEN purposes we propose the following definition:

> *A forest user group (FUG) is a formal or informal association of forest users, responsible for a set of activities related to setting, implementing and/or enforcing rules of forest use and management. The group must have a minimum level of organization, including a clear understanding of who are the group members, regular meetings at least once a year, and some joint activities during a year (but all members may not necessarily participate in these).*

4.6 Village

Using villages with very different sizes can cause some biases when it comes to comparing them. One may, for example, incorrectly conclude that one village has better infrastructure (e.g., access to credit institutions) just because it is larger. If village comparison is an explicit goal in the research, selection of similarly sized villages becomes an issue.

For some countries the term ‘village’ does not exist, and ‘community’ or similar might be the preferred term to be used during the survey. For purposes of PEN research a village is defined as the lowest administrative unit in an area, normally under the jurisdiction of a village leader/council. In some countries the village can be very large, several thousand households. If the “villages” typically have more than 500-1000 households one should consider breaking it up into smaller units and sample one or a few of them.

The PEN researchers should consult the PEN coordinator or advisor if in doubt on the relevant village definition for the study area.

4.7 Household

A household is defined as a group of people (normally family members) living under the same roof, and pooling resources (labour and income). Labour pooling means that household members exchange labour time without any payment, e.g., on the farm. Income pooling means that they “eat from the same pot”, although some income may be kept by the household member who earns it. One should also note that it is possible to have household members who are no blood relatives of the family, e.g., a household servant, an in-law, or someone taken into the household because they have been orphaned or otherwise destitute.

In most cases the members of the household are easily identified, but here are some difficult cases:

1. Polygamy: If a man has several wives, each living in separate houses, then each of the houses of the wives should be treated as a separate household. The man's contribution to the particular household selected should be included. (Note: if the sampling is based on selecting household head, and a man with two households is included in the sample, one should include both households and register them as two separate ones.)
2. Several families living together in one house: If there is resource pooling, they should be treated as one household, otherwise not. An example: A married son still lives in the house of the parents. If his family's economy is separated from the one of his parents (no income pooling, and they only occasionally eat together), they should be treated as two separate households.
3. Family members living away parts of the time, e.g., working or going to schools: Include if they are more or less fully integrated in the household economy (e.g., school children living away during the week, but parents paying for their expenses). In other cases were children have left the house to work and take care of themselves, but contribute some income to the remaining household, they should *not* be included in the household. But, the money contributed should be recorded under remittances.
4. In some situations an extended family may be living in different houses, but sharing the same land and income, and eat most of their meals together. Again, following the household definition of "resource pooling", the extended family should be treated as one household, although they don't sleep under the same roof.
5. One may have single person households, e.g., a widow living alone.

Household head

The household head is normally well defined by customary and/or official rules. In most cases this would be the husband, but this might be more complicated if several generations are living together in the same household. But, simply asking "who is the head of the household?" would normally solve the problem. In some contexts the customary head is not the same as the official one, in which case one may choose to follow the customary rules.

5 A detailed guide to the PEN prototype questionnaire

This section provides a guide to the PEN prototype questionnaire. Together with the definitions and code list this should enable PEN researchers and enumerators to implement the questionnaire in a uniform way, which is needed to generate comparable data for the global analysis.

5.1 General rules and hints

Fill in common products in tables!

In the quarterly surveys, there are a number of tables with quantities, prices, etc. as columns, while different products constitute the rows. These rows are empty in the prototype questionnaire. The PEN researchers should fill in the most important products in the actual questionnaire used, based on information gathered during PRA and pre-testing of the questionnaire. It is critical that such a list is made and included in the questionnaire used, as respondents will not remember all the products collected or grown unless asked specifically about them. We suggest the following procedure:

1. Get a gross list of agricultural and forest products at the national level (official statistics, research reports, etc.)

2. Use the RRA/PRAs, the first village interview (V1), questionnaire pretesting, and proper observation to consolidate the relevance of that list at the district/village levels.
3. **Include (print) the list of village-relevant products directly into all the household questionnaires** (no messing around with separate lists!), so that the enumerator will remember to ask for all the potential products the household cultivate or harvest.

At the end of each table one should also ask if any other products not mentioned are being collected/harvested or cultivated/produced. Remember to leave space for additional products.

For other items the various products have been filled in, e.g., in section I on livestock (Q1-Q4). One may leave out items not relevant for the local context to shorten the questionnaire and save printing costs (but remember to keep the codes, don't renumber them).

Getting the time frame right

All the income questions (and some others) refer to a particular time frame, i.e., past month (30 days), past 3 months (90 days) or past 12 months (365 days). Make sure that the relevant time frame is clearly understood by the respondent.

Also note that a phrase like 'past 3 months' means the '90 days prior to this interview', and *not* – if the survey is done in mid-August – the months of May, June and July. Beware that, unless you make your instructions very clear, some respondents will answer in terms of their lifetime experience, others will answer in terms of today, and everything in between.

Ranking

Several questions in the village and annual household survey require ranking of alternative responses ("rank the most important, max 3"), e.g., the reasons why the household joined a forest user group (FUG). These questions can be asked in two ways: (1) the enumerator can read the question, then read out all the choices first, and finally ask the respondents to rank these from 1-3 (and leave out the rest); or (2) the enumerator read the question, and then allow the respondent to answer the question without guidance from the list. The enumerator simply marks the choices that happen to be answered.

PEN researchers should use the second method. Thus the ranking questions should be asked in the following way:

1. The enumerator asks the question as an open ended question. Example: "What are your reasons for not joining the FUG?" (A1F11)
2. The respondent gives one or more reasons, e.g.: "I'm busy working in the fields". The enumerator should relate this answer to the choices given in the questionnaire, in this case option 4: "Cannot afford to contribute the time".
3. Follow-up questions can be asked to clarify the response, or to ask about other reasons if they only give one reason.
4. The enumerator asks the respondent to rank the reasons given: "Among the reasons given, which one is the most important?" And then: "the second most important", and "third most important" (if applicable).

Note:

- You don't always have to fill in for all 3 alternatives, i.e., in some cases you only fill in a 1, or a 1 and 2. Don't force the respondents to give 3 alternatives.
- The answers given might not correspond completely with the wording of responses given in the questionnaire, but some flexibility is allowed, provided the meaning is the same. But,

obviously if it doesn't fit, you should use the "other" category. Where to exactly draw the line is hard to tell and it is up to your good judgement! Determine the guidelines as best you can on the basis of your pretesting experience. Encourage the enumerators to write some explanatory text whenever they use the 'other' category (you may regroup it during coding and data entry).

- You should only fill in 1, 2 and 3 for the ranked alternatives, and leave the other cells blank.
- What does "most important" mean? In some cases this can be interpreted as "highest economic value" or "highest impact on household welfare and livelihood". But for many questions, like the one used as example above, it is simply the subjective assessment by the respondent.

Cells that can be calculated

Several cells in the quarterly survey can be calculated, and these have a light grey background. The commonly calculated cells are:

- Production or collection = own use (household consumption and gifts) + sold (in markets and for barter).
- Gross value = production (collected) * price per unit.
- Net income = gross value – purchased inputs.

You don't need to calculate these during the interviews, nor during data entry. In fact, in most cases it would be a waste of scarce time to do so, as this can be cone when checking and coding the questionnaire, or when entering the data.

Note:

- Sometimes you may want to use the totals as an explicit way to check the consistency of responses. The interviewer should be the one who quickly makes the calculus, and presents it to the respondent for reaction, and possibly a revision of the first answers.
- In some cases the respondent may not know the disaggregated variables, but only the product or sum. For example, they may know for how much the sold of forest product X, but not the quantity and unit price. In this case you should write down the total, and then fill out the quantity and unit price based on general information collected on the unit price of the product. In other cases, such as for the distribution between own use and sold, one should try to get a rough estimate from the respondent. Again, make notes during the interview to help in later coding and data entry.

Using local units

Local units of measurement (rather than metric ones) can be used in filling out the questionnaire, provided they do not vary too much across households. An accompanying list of conversion factors should, however, be made and accompany the PEN data set to be submitted. See section 9 of the appendix 1. The conversion factors should be generated based on a small survey of, say, 10 households. Note that there may be large local variations in the meaning of 'headload' or 'bag', so getting the conversion factors from several measurements is important. Then use an average (mean) of these, possibly after excluding some outliers.

Also note that there is a big difference in the weight/volume ratio between fresh and dried products, e.g., wood and weaving materials. Household A may refer to x kg of dried material, while household B refers to y kg of fresh material, and the answers are therefore not readily comparable. You should then convert everything into either fresh or dry (whatever you chose as the default) to get consistency across households. In other words, you may ask the households for the quantity of either dried or fresh weights, but then use a conversion factor between them, such

that what is in the end entered in the database refers to your default choice (weight of either dried **or** fresh material).

Illegal activities and underreporting

Some forest products are harvested illegally, although the border line sometimes is difficult to draw. This may lead to serious underestimation of important forest uses, such as timber and game meat. There are no shortcuts to getting the true answers, but the following hints may be useful:

- **Patience** and building **trust** are keys to get correct information on illegal activities.
- Establish a reputation as a researcher wanting to find out what people do, not someone who is there to conserve the forest or on similar missions. PEN partners are not in the field to judge the morality of villagers' activities, just to record them. It may be worth stressing repeatedly the neutrality of your research in this respect.
- Some have also suggested that you should actively show that you accept illegal activities, e.g., buying and eating illegal game meat. But, there is a line not to be crossed here, and we leave it to the individual partner to decide on the location of that line.
- Related to this, avoid being connected to government officials or NGOs in the area, which can make respondents reluctant to reveal the true figures. Never conduct the interviews together with officials, neighbours or other non-household members present.
- If you suspect that some households underreport on their forest use, it may be worth approaching them in an informal setting (no pen and paper), and then record the information immediately after or return to them at a later stage to try to get the details. Spend more time on the "difficult household", and possibly visit them yourself (not the enumerator).

Codes

The codes to be used are included in Appendix 1. Getting uniform coding is extremely important for the quality of the data in the CDS. We realize that the list of codes is not exhaustive (e.g., for crops), and there is space for adding new codes. Suggestions for new codes should be sent to the PEN coordinator, who will ensure that an updated list of codes is available on the PEN Web site.

You will have to make the codes for district, village, household and personal ID yourself.

District identification number (DID): number the districts consecutively (01, 02, ...). The district code can be useful for later categorisation of villages.

Village identification number (VID): The simplest is to just number the village consecutively 01, 02, 03, …. Alternatively, if you work in several areas/districts, and want to categorize households or villages according to these in your later analysis, you may start label village 1 in district 1 the code 11, village 2 code 12, and then village 1 in district 2 gets code 21, and so on.

Household identification number (HID): The households in the sample are given a 3 digit code. These are numbered consecutively from 001 for the individual PEN survey (i.e., they don't start at 1 for each new village). Alternatively, you may prefer to give the households in village 1 numbers from 100 to 1xx, in village 2 from 200 to 2xx, etc.

Personal identification number (PID): Each person in the household is given a 2 digit PID. It is used later for wage income in section F in Q1-Q4.

Remember to keep an up-to-date list of the DID, VID and HID.

Your PEN partner code is given on the PEN Web page (under partners).

5.2 Getting information in the village questionnaires

The information in the village questionnaire can be obtained from a variety of sources, depending on the local context and the specific question. Here are the most common sources of information:

1. **Own observation/measurement:** Some information can be obtained by the PEN partners or enumerators own observation or measurement, e.g., "distance to the nearest road usable during all seasons". One may also use GPS to measure distances.
2. **Secondary data:** Some villages may have good records of population (time series), access to public services, land categories etc. that should be used in filling out the questionnaire.
3. **Village officials:** In cases where reliable written records are not available, village official may have some of the factual information needed. **Warning:** Don't rely (only) on the official version for answering questions involving some assessment (sections E to G in V1) – it may give a biased and "polished" view of the state of affairs.
4. **Key informants:** In most villages one will find people that have a good overview of the situation, and these can be very useful in providing inputs. But again, cross checking information is vital as there is always a degree of subjectivity and uncertainty involved.
5. **Village/focus group meeting:** If at all possible, it is **strongly recommended** to have at least one focus group or village meeting in every village to go through the qualitative questions of sections E-G in V1, and possibly also some of the factual ones if they cannot be answered from any of the above sources.
6. **Village census:** A village census, to be done before the household survey starts, can provide key demographic information for the village questionnaire. See section 6.3.

A village or focus group meeting can be organized in different ways. Depending on the size, one alternative may be to invite all village members. Another option is to call a smaller group, say 8-10 people. In many countries it is expected that you to go through the village leadership when organizing such meetings, not doing so can be seen as both impolite and possibly also a direct violation of the rules and regulations and can obstruct the research seriously. But, you should also be aware that a village leader may select an unrepresentative group. You should therefore stress that you would like to get a diverse (men and women, young and old, rich and poor, immigrants and long time residents, etc.).

During the village/focus group meeting: have the list of questions ready and ask them in a systematic way. Follow up on interesting leads, but don't get side-tracked. Many prefer the semi-structured form, and the village questionnaire provides such a structure. Fill in as much as possible into the village questionnaire on the spot.

One may provide a small snack/drink during the meeting, which should not last for more than two hours.

If contradicting views and information occur during the meetings, try to reach a consensus answer. More generally, it is critical to **double check information** given by individuals. Thus, ask the same question to many individuals (even though the translator or research assistant may think you have a bad memory). This is particularly important for information that potentially could be sensitive, controversial or particularly important for your own key research questions.

The PEN partner should be present during village meetings, and be responsible for filling in the village questionnaires – this task cannot be delegated. This is unlike the household questionnaire where, after an initial training period, the enumerators will do most of the data collection. The reason is that the information in the village questionnaires requires a more critical assessment and judgement.

5.3 Village Survey 1 (V1)

It would make good sense to do the village survey and fill in the questionnaire at the beginning of the fieldwork (as planned), but then go over the survey at the end and check and correct answers based on what is learned during the fieldwork.

The PEN partner should review the definition of a village in section 4.6, and consult the PEN coordinator if the concept is not clear in your setting.

A. Geographic and climate variables

- Give each village in the survey an identification number (VID), see end of section 5.1.
- Latitude and longitude data may either be available from GPS, or from secondary sources. Everyone should therefore be able to fill in these (Q3-Q4).
- Rainfall information should be obtained from the lowest administrative unit for which this information is available (or nearest meteorological station).
- If data are not available for 20 years, use shorter time period (and specify in village narrative).
- The coefficient of variation (CV) is the standard deviation (SD) divided by the mean (M): CV = SD/M.

B. Demographics

- See definition of village in chapter 4.6. Also note that the term "village" refers to the physical area. In some cases people may have moved "the village" from elsewhere. The year of the establishment then refers to when the village was moved, i.e., year of establishment in its present location.
- Exact figures are not expected for the in- and out-migration, but try to give a rough estimate. Sometimes it may be better to ask about proportions, and then multiply with population to get the answer to be filled in.
- Much of the information in this section can be collected through a village census, to be done before the surveys start (if done at all). See section 6.3.

C. Infrastructure

- Again, it may be better to ask about proportions rather than absolute numbers on access to electricity, piped water etc., and then multiply with number of households. These variables will be used as indicators for the level of development in the villages.
- Q6: By this question we mean if it is possible to reach the village by road during the entire year (i.e., not roads inside the village).
- Q10: the district market refers to the larger market, and would normally be the market located the furthest away from the village among the four listed, except in the cases where it coincides with the markets mentioned under 2-4.

D. Forest and land cover/use

- Recheck the definitions of land cover/use in chapter 4.1 - 4.2.
- The question is far easier to answer with a local map at hand. A map based on remote sensing would naturally be the preferred option, if available. If no map exists, you may sit down with a few informants and make a rough map of the village land. With a map you can make some rough estimates of the size of different land categories if exact information is not available.

- The first table in this section contains a simplified description of ownership, and the following definitions apply. The first three ownership categories (state, communal and private) refer to the formal (*de jure*) ownership categories, with some enforcement of rules of access, uses and management, cf. *code-tenure* in Appendix 1. The last category (open access) refers to land without any enforcement of rules. Thus, we have the following categories in the table:
    - **State ownership**: the formal owner is the state at national or regional level, and there is some degree of rules enforcement.
    - **Community ownership**: the formal owner is the community, and there is some degree of rules enforcement.
    - **Private ownership**: the formal owner is private entities (individuals or companies), and there is some degree of rules enforcement.
    - **Open access:** the formal owner is either the state, community or private entities (and in a few cases no formal owner), and there is no enforcement of rules of access and use.
- Table 2 should be consistent with table 1 in terms of ownership categories (more detailed in table 2) and total area.
- Q3 on management easily invites villagers to give you the "sunshine story". Double check by being more concrete, for example, by asking if they have done this specific management task over the past year.

E. Forest resource base

- As noted in the questionnaire, this is a question suitable for a village/focus group meeting. One would typically ask the question in an open form, and let the discussion flow while you take notes of the answers. Then, for the ranking questions, you ask them to state which of the many reasons mentioned is the most important, and then the second most important, etc.
- Take note of the two footnotes below the table!

F. Forest Institutions

- These questions should be asked in the same way as section E above.
- The information needs an active assessment by the PEN partner, and a critical attitude is needed. For example, people in the village may **say** they respect government rules, but they are *de facto* not. Thus if you get information during key informant interviews that government rules are commonly disregarded, you may trust this more than the official version given in a village meeting. In some cases, it may be appropriate to confront the meeting participants with this information to double check (but this depends on the sensitivity of the issue and local culture, use your discretion). Concrete follow up questions are also helpful to get closer to the true picture. For example, you may ask: "do some people collect firewood for sale from the forest reserves?", knowing that this is illegal.

G. Forest User Groups (FUG)

- Forest User Groups are not found everywhere (e.g., rare in Latin America). If they don't exist in the area, you should omit this section from the questionnaire used in the field, and inform the PEN coordinator and advisor about this decision.
- The information in this section would typically be sought from members of the leadership of the FUG. Get 2-3 people together to answer the questions, but you may also want to double check some information with regular members.

- Q9 on development projects should include any development efforts using money generated by the FUG, e.g., a new water pump, a new road, or rehabilitating the health centre.

5.4 Village Survey 2 (V2)

A. Geographic and climate variables

- Q3: If exact rainfall data are not available, use a subjective assessment after talking with a few key informants.

B. Risk

- The purpose of this question is to find out if the past year (covered by the household surveys) was an unusual one, and if so, in which respects. To many poor people life may be a permanent crisis, and this may lead to overstatements compared with what is intended with this section. For example, if a certain crop pest is common (occurs every year), and has not been particularly severe this year, it should not be included as a crisis.
- Three codes are possible: 0=no crisis, 1=moderate crisis; 2=severe crisis. In general, one should be careful in using category 2, and let that refer to more catastrophic events that leads to widespread sufferings and possibly the loss of human lives. As a rule of thumb, a crisis of category 2 should be one the leads to a drop in the overall village income by more than 1/3.

C. Wages and prices

- Q1: In case of no seasonal variation (peak and slack season) in the wages and prices, one should just fill in the same figures in both rows.
- Q1: In some cases the wage may include some in-kind payment, typically food (either a free meal or some staple food to take home). The value of such in-kind payment should be converted into a monetary value (e.g., by asking how much the wage rate would hypothetically be if the meals were not included) and included in the wage rate reported. Thus the wage should be both in cash and in-kind.
- Q4: In some villages land sales may be prohibited or so uncommon that the land value cannot be estimated. Use then the code "does not apply" (-8).
- Q4: The land value should be for land itself, and should *not* include any extra work done (land improvements), trees planted, or the value of any buildings on the land.

D. Forest services

- Q2: Note that only the payment to the villages should be included here (even if not all the households in the village will benefit equally from the payment). Payments directly to individuals/households may be identified here, but the quantities should be included in A2 instead.

5.5 Annual household survey 1 (A1)

A. Identification

- For the identification numbers – DID, VID, HID and PID, see end of section 5.1.
- Q6: Not all studies will have access to a GPS, and the question is therefore not obligatory. A GPS is, however, recommended for two reasons: measurement of distances and areas, and identifying households for later surveys (later research projects, but even for the next quarterly survey).

- Q7: The same central point should be used for all households in a village.

B. Household composition

- The names of the household members will not be used later on in the data analysis, but it might be useful to have the complete list: (1) some questions later (e.g. wage work) relate to household members and can then refer to names; (2) if it turns out there is an error in the enumeration concerning a particular person, it is possible get the correct information during a return visit to the household; and (3) when returning to the household, it would be an "ice-breaker" to remember the names of the household members (have a quick look before arrivial).
- Table 1, column 3: Instead of year born on may ask about *age*, and then recalculate when entering the data. Also, one may ask about year born according to national calendars (e.g., in Ethiopia and Nepal), and then convert to standard the standard international (Gregorian) calendar.
- Table 2, Q1: Couples living together without being married (cohabiting) should be coded as married. In the same way, "spouse" (code 1 in Table 1) also covers those cohabiting in a "non-formalized marriage-like relationship".
- Table 2, Q2: The time of household formation would in many cases be the time of marriage, but not always. In some cases the just married couple may live together with the parents of either the husband or wife and practice "resource pooling", before they eventually get their own house and household. Given the definition of a household (4.7), the year of formation will then be when they move out of the parents' house and started living on their own. But, admittedly, there can be more complicated situations, and these should be discussed with your PEN advisor/coordinator.

C. Land

- Refer to the definitions of land cover/uses (sections 4.1-4.2). Make sure these are clearly understood by yourself and the enumerators.
- Drawing the exact line of how much land a household owns can at times be difficult. For example, in a shifting cultivation system the family lineage may have customary rights to forest that has been cleared for shifting cultivation but is currently natural forest. But, these rights tend to weaken if many years have passed since the last cultivation. One critical question to figure out ownership is to ask: "Can the household deny others in the village from using that land for cultivation?" Thus ownership refers to the households' exclusive rights to the primary uses of that land.
- Household may have several plots in the same category, e.g., cropland, and the plots may have different ownership status. One should then fill out the *dominant* ownership category (in terms of area).
- Some land, for example, pastures, may be shared with other households. Try to make a reasonable estimate of the respondent's share (and record that) based on, for example, the relative number of cattle grazing or the number of families sharing it.
- Some households do not know the size of their plots, others deliberately understate the size (particularly if taxation is based on land holding).
- It is often recommended that the enumerators go out and measure the fields. This is most easily done with GPS (if the plot is not too small), or by some manual measure tools (using a rope of 50 m, with labels every 5 meters will do fine).

- The ranking of the main products refer to their "overall contribution to the wellbeing of the household over the last 12 months". The simplest measure would be the total income from that crop, that is, both subsistence use and cash income. Beware that some respondents (and enumerators!) may think in terms of cash income only.
- In the ranking of the main crops, grass or legumes for grazing domestic animals are options that should be considered (see codes 421-). The ranking here should reflect the approx. value of the grass/legumes rather than that of the animals (which require many other inputs).
- If several couples are living together in the same household, one person might not be aware of the size of other person's plot. It will be more accurate to ask each person about the size of his/her plot(s).
- After 2-3 rounds of the quarterly survey, it may be useful to revisit this section to check if the information given makes good sense and is consistent with, for example, the agricultural output reported.

D. Assets and savings

- The purpose of this section is to get a rough indication of the wealth of the household, which can be used in later analysis of, for example, how forest dependence is varying according to household wealth. To be able to aggregate into one unit, it is therefore important that the current value of the items is indicated (i.e., not the value of the item when the household acquired it).
- Q1: The type of house is often a good indicator of wealth. If people have several houses far apart, get information for the house in which they normally live. If the house consists of several unconnected buildings close to each other (e.g., a separate kitchen unit), the total size (m^2) should be recorded.
- Q2: Respondents may tend to give the purchasing price of the items rather than the current (sales) value. One may then also ask about the age, and write off the value accordingly: current value = new price * (age/expected life-time). Alternatively, one may choose write off a fixed depreciation percentage annually (e.g., 10%). Whatever method is chosen, it is important to be consistent across all households in the survey.
- Questions about savings and debt are often sensitive, and may therefore better be asked at the end of the interview.

E. Forest resource base

- Q3: It's difficult to got good time estimates; people don't wear an arm watch, and firewood collection is often combined with other tasks, e.g., they bring firewood when coming back from the agricultural field. The suggestion is that when collection is done as part of such multi-tasking, one counts half of the time spent as time for firewood collection.
- Q5: Note that the question refers to the physical availability of the forest resources, and not the actual use by the household.

F. Forest User Groups (FUG)

- One should first, through RRA and V1, establish if there is any FUG in the area, and then ask explicitly if the household or any member of the household is a member in any of them.
- Q4: Thus question may be too complex to answer by the respondent, and one way to do it would be as follows:

a. Asking and listing the types of activities

b. Asking and filling the frequency/year of each activity

c. Asking the number of household members involved in each activity

d. Asking the number of days spent on each activity

e. The total person days can be calculated by multiplying and summing for each activity

- Q5: By "contributions" is in-kind (non-cash) inputs, e.g., seedlings and hives. Try to estimate the value of this and include that value together with any cash contributions. Don't include the value of family labour contributions, that is covered in Q4.

5.6 Annual household survey 2 (A2)

A. Identification

- See A1.

B. Crisis and unexpected expenditures

- As for V2-B, the purpose of this question is to find out if the past year was an unusual one for the household, and if so in which respects. The same warning as at the village level about an overstatement of the crisis also applies here.
- Three codes are possible: 0=no crisis, 1=moderate crisis, 2=severe crisis. Code 2 should refer to events that lead to a *significant decline* in the household well-being. To get uniformity across the studies, the following rules of thumb may be used for code 1 events (moderate crisis): crop failure within 1/3 of normal range, illness for less than 3 months, losing only a small proportion of land, cattle or other assets, and losing a low-paid part time job. But, a household-specific assessment is often required, and the good judgement of the PEN partner and enumerators is needed: the death of the main "bread winner" within the household may have large well-being consequences, and should be a category 2 crisis, whereas the death of a newborn baby may be a category 0 or 1 crisis in terms of its economic impact.

C. Forest services

- Note that many respondents know what the various terms mean, so explain that. The exact grouping may also be necessary to do yourself based on your knowledge of projects in the area, e.g., to distinguish a "carbon project" from one of "biodiversity conservation".
- Some projects may be payment for more than one service, for example, *both* for carbon sequestration *and* for biodiversity conservation. If possible, the amount should be split based on their estimated shares. The share can be estimated both based on the respondent's answer and information from the project or village officials. The amount received should in any case NOT be written twice, i.e., under both carbon and biodiversity, as these will later be added up.

D. Forest clearing

- Q9-Q10: Remember that the information here should include the clearing over the past 12 months, i.e., the amount cleared last 12 months (Q2) should be included in the amount last 5
- Q11: This figure should include land left for fallow. It should also include situations with light enrichment planting of the fallow, e.g., by rattan, rubber, or fruit trees.

E. Welfare perceptions and social capital

- The purpose of this section is to get some non-material and subjective indicators of poverty and well-being, to ask about perceptions of change, and get some brief information about social capital (trust, reciprocity). This can be used to see how, for example, forest dependence relates to subjective well-being.
- Q1: This question attempts to cover what in the literature is referred to as the 'quality of life' or 'subjective well-being', i.e., both material and non-material aspects of life. Another word describing this is '*contentedness'*. This word requires careful translation into the local language. Avoid words that refer to the material aspects (income and assets) only; that is for the next question. Also try to avoid words referring only to 'happiness', which most people tend to say they are (and therefore is less useful for other purposes).
- Use the full scale, and be careful not to cluster answers in the middle.
- Q2: Ask the question in such a way that income is not just referring to cash income.
- Q3-5: The term 'well-off' refers to material aspects, i.e., income and assets. One may use terms equivalent to that, or refer to 'poverty' (although the meaning of 'poverty' varies across cultures).
- The ordering of the questions is important (as have been shown in many surveys), and the listed order should be followed during all interviews. For example, if Q2 asked before Q1, then Q1 gets a more materialistic interpretation than intended. Similarly, Q2 should get the assessment of own needs vs. actual income, and not compared with what others have; that's for Q3.

F. Enumerator/researcher assessment of the household

- This section is filled out by the enumerator and/or the researcher (PEN partner), and should not be asked or discussed with the respondent. The first question must be answered by the person (enumerator or PEN partner) doing the last interview, while the others should be filled out by those who have had most exposure to the households (done most of the interviews).
- The first question about smiling and laughter may appear strange, but there is a significant literature on how smiles correlate with happiness and well-being (including work done by some of the resource persons of PEN). Thus with a simple, smiley question one can get 1-2 more research papers out of PEN. And, possibly add some smiling to the life of the enumerators.
- The information about survey quality can also be used to test potential biases, e.g., if there is systematic underreporting of forest income.

5.7 Quarterly household survey (Q1-Q4)

Recall periods

- It is critical to be fully aware of the variable recall (recording) period for the income questions in the quarterly survey. For the first sections (B-G) covering income from forests, fishing, other environments, wage and own business, the recall period is **one month**, meaning the 30 days prior to the interview. For income from agriculture and other sources (sections H-J) the recall period is 3 months, meaning the 90 days prior to the interview. Thus, when moving from section G to section H, make it absolutely clear to the respondent that the recall period is changing.
- The reason for the difference in recall period is as follows. Generally, the shorter the recall period, the more accurate the data. However, most agricultural income is in the form of 1-2 major harvests, and using a one month recall period in a quarterly survey may miss the harvest and therefore be very misleading. A similar argument applies for remittances.
- For informants without a good sense of the calendar, using special events (village festivity, big rain, birthday of a member of the household, etc.) will help them to identify the recall period.

- Using one months recall for forest income data in a quarterly survey causes one potentially significant problem: Some forest products are collected for very short periods of time, maybe only a couple of weeks, yet they are of major importance (e.g., mushrooms in Zambia, or Brazil nuts in Brazil). The quarterly survey can be "lucky" to capture that, i.e., the collection happened during the 30 days prior to the survey, or it may be "unlucky" and miss it (happened 31-90 days prior to the survey. PEN recommends the following procedure to deal with such cases:
    - During RRA, V1, pretesting and other sources of information, one should identify important forest products that are harvested very irregularly and/or only for a very short period (less than a month) of the year.
    - For these products one can ask, after the normal questioning of forest products and one month recall, if one has collected any of these products **during the past 3 months**. If they have, the quantities of these products should be divided by 3 and then entered the table in the same way as other forest products. Why divide by 3? The forest income data will be multiplied by 3 to get quarterly income, and then aggregated to annual income. Thus if not dividing by 3 the quarterly and annual incomes will be incorrect.
- The use of 3 month recall periods in a quarterly survey means that the timing of the surveys is important – it should be 3 months between them! While a few days deviation is acceptable, having several weeks deviation will lead to either overlapping survey periods and double counting (if shorter), or to gaps (if more than 3 months). If you for reasons beyond your control the timing between the surveys deviates with more than 2 weeks, this should be discussed with the PEN coordinator.
- The above also implies that the implementation of the quarterly surveys should be done in the same order, i.e., the first village interviewed for Q1 should also be the first village for Q2, and so on.
- For Q2-Q4, the question about "past 3 months" can in some cases be replaced by "since the last survey", particularly when you have the previous survey at hand. The memory of the respondents is also imperfect. For agricultural income one therefore needs to look at the previous quarterly survey to make sure that there neither double-counting nor harvests being left out. For example, you want to bring along the previous surveys, and ask "do you remember in a previous survey you told me you have a 0.5 hectare plot with maize? What happened to that maize?"
- More generally, it is useful for the enumerators to go through the previous surveys to get "context" information before stepping into the household with a new survey. But, watch out for enumerators (and respondents) just repeating the information provided in the previous survey. It's a simple shortcut, but not acceptable!

Storage

- Does forest, environmental and agricultural income over the past 1 or 3 month(s) refer to the value of what is collected/harvested, or what is consumed and sold over that period? The short answer is that the *collected/harvested* amount is what we want to know. Thus if 500 kg of rice is harvested, but 300 kg still stored at the time of the interview, 500 kg (and not 200) should be recorded as the production, and the *planned* use of the 300 kg stores should be filled in (i.e., for own use of sale). This may of course cause some inaccuracies (don't know planned use, or the actual use may be different), but the alternative ways of recording will be more demanding. One optional more time demanding – and accurate – way of addressing this is to keep the space empty and fill the information when visiting the household for the next quarterly survey.

- The above principle also applies to wage income: the amount earned in the period should be recorded, even though the cash payment may not yet have been received (or received in advance). Therefore, if people receive the money during the next household survey, it should not be recorded for a second time. For remittances, gifts and other unilateral transfers (section J), however, the time of receiving the money is the relevant one as there is no "earning period" as such.

Other general issues

- The same product may be collected from different types of land, sold in different markets, etc. In other words, the information asked for in, for example section B, is not uniform for a particular product. There are then two options: (1) If there is one dominant market for the product, this market should be filled in and only one row used. (2) If this cannot be done, one can fill in several rows in the table for the same product. For example, the first row states the sale of the product to consumers within the village (code 11), while the second row states the sale to a private wholesale buyer outside the village (code 22).
- It may be useful to make a village price survey for the most important products, and use that as a guideline when filling out the household questionnaires. However, the prices can vary substantially between households, depending on, among other things: (1) *Season:* lower prices just after harvest. (2) *Where and to whom sold*, e.g., bringing the products to the marketplace may give higher prices, but also take longer time. Markets may also be interlinked, e.g., a product is sold to a trader that also provides credit, and therefore offers a lower price. (3) *Quality differences.* Thus variation in the prices obtained varies across households, and this should be reflected in the questionnaires. But, one should also watch out for very high/low prices; maybe the information provided is wrong (or the unit used is not what it is said to be). In short, the typical village market prices should be used:
  (1) for rechecking with the household suspiciously high/low prices reported by the household, e.g., more than 50 % above or below the village price;
  (2) as estimates of the price the household receive in cases where they cannot remember the price or the total value of what they sold (but only the approx. quantity).
- Forest products are often collected from a limited set of forest areas. It may be useful to make a list of these, and just fill in the name (or a preliminary code) for these during the interview, with the proper PEN codes being filled in later. This will both save time and ensure consistency among the enumerators.

A. Identification

- See A1-A.

B. Direct forest income (income from unprocessed forest products)

- As noted in the code list, an important distinction is made between unprocessed (raw-material) forest products and processed forest products. "Processed" means a significant modification or change of the product, e.g., turning wood into charcoal or a chair, or turning clay into a pot. Minor modifications, for example, cutting rattan canes or bark into smaller pieces, or washing and drying the product would not qualify, and the products should still be classified as unprocessed. If in doubt, you should consult the PEN coordinator and/or advisor.
- Keep in mind the forest definition of PEN, which includes plantations. This means that income from plantations, including woodlots larger than 0.5 ha, should be included in this table. Income from smaller woodlots should be included as agricultural income (section H).

- Also keep in mind that agroforestry income is not classified as forest income, but is included under agricultural income (section H).
- If the product is not sold in a market, you may fill in column 10 (type of market) fill in code -8 for 'does not apply'.

C. Forest-derived income (income from processed forest products)

- See distinction above between unprocessed and processed forest products (5.7.B).
- This section asks in the first table about the quantities produced, prices and non-forest inputs. Then the second table asks for the forest inputs (raw materials) used in producing the products listed in table 1.
- Avoid double counting: unprocessed forest products collected for use in the household production of forest products should be included in table 2 of section C, and **not** in the table of section B. Only products collected (and used by the household or sold) without being processed by the household should be included in section B.

D. Fishing and aquaculture

- The land type (columns 2) refers to the dominant land type surrounding lakes or rivers. The ownership (column 3) will then refer to the dominant ownership of that land.

E. Non-forest environmental income

- This section covers all types of environmental income, except for forest and fish related which is already covered in sections B – D.
- For game meat, the land type (column 2) refers to the main living habitat of the game, rather than where it is caught. Information of that could be based on information from key informants.

F. Wage income

- If the payment is (partly) in kind (e.g., helping in harvesting and get paid 10 kg of rice), you should estimate the monetary value of that and include as regular pay in the questionnaire.

G. Income from other own business (not forest or agriculture)

- By "own business" is meant that a person(s) in the household owns the business. In economic terms, ownership is usually defined as being the "residual claimant", i.e., the owner has the right to the profit generated (what is left after all the expenses is deducted from the sales income). This could, for instance, be running a small kiosk, a barber shop, renting out horses, or using the car or motorcycle as a taxi.
- If the business is shared among several persons (households), the share to the members of the responding household should be reported (both for gross income and costs).
- Income from forest-based business should be included in section C, under processed forest products. The same is true for agriculture (including livestock): a small family business which produces, for example, cheese or ghee, should be reported in section I.
- However, if the forestry and/or agricultural business is run on a larger scale, i.e., it involves hiring non- family labour on a regular basis, and the products are for market sale, then the business should be recorded in this section.

- Q4: own non-labour inputs refer to, for example, agricultural products from the farm being used as inputs in the business. The equivalent market value should then be stated.
- Q10: this question is included to estimate the costs of having capital (depreciation & opportunity costs of capital), which should be deducted when estimating net business income. Note that there may be some overlap with the household assets recorded in A1D2, but that is fine (double-counting is ok in this case).

H. Income from agriculture – crops

- It is important to note that this section includes both annual and perennial crops. Thus **agro-forestry** as well as trees-on-farm outputs (e.g., horticulture tree products) and inputs are recorded here. Also crops grown in (home) gardens should be included here. In short, **all crops** grown on land not classified as forest (plantations) should be included in this section.
- Valuable agricultural by-products should also be included.
- Table 1, column 2: The area of production means the total area used for that crop. In the case of intercropping, you should get a rough estimate of the *share devoted for each particular crop.* In other words, when adding up land area for each crop the total should correspond to the total land the household cultivates, as also reported in A1, section C.
- Land that is rented in (borrowed) falls in three categories which are treated as follows:
    1. The rental is paid for in kind by giving the owner a share (*sharecropping*) or a certain amount of the output: the production figures should state what the household in the survey (the borrower) keep, and not the total production. (The income to the landowner – if he/she is included in the survey - should be given in section J, Q6.)
    2. The rental is paid for in *cash*: the costs should be stated under "inputs" (table 2, Q20). (And the income to the landowner also recorded in section J, Q6.)
    3. *No rent* is paid: the table is filled out as if the borrower household owned it. There are no economic transaction in relation to the rental to be recorded

    In any case, remember to include rented in land in A, section C, Q12.
- For inputs, the key is to get the total costs in order to calculate net agricultural income. Thus, the farmer may, for example, use different types of seeds, but this is only relevant in order to estimate the total costs of seeds. If you prefer it is, however, possible to enter different types of seeds (or other inputs) several time (use same input code, but different quantities and prices).

I. Income from livestock

- The questions about livestock in table 1 may work better if you ask in either of two ways:
  (1) start by asking with how many heads they have now (column 7), then ask how many they had 3 months ago (column 1), and then ask about column 2-6 to explain the change over the last three months; or
  (2) start by asking with how many heads they have now (column 7), then ask about any changes over the past 3 months, and finally end up with the number 3 months ago.
- In table 1, make sure that there is consistency between the surveys. For example, column 7 in Q1 should be the same as column 1 in Q2.
- In table 2, information about meat consumption and sale is already included in table 1. Thus you may not need to ask it again, but have to ask about the price.
- In table 2, make sure that milk production does NOT include milk used for production of butter, cheese, ghee or other milk-based products. This would lead to double counting and

overestimation of livestock income. Thus, only milk that is consumed or sold as milk should be included. (Similarly for eggs: only eggs consumed or sold should be included, not those hatched.)

- In table 2, draught power: *Own use* refers to how much, for example, oxen are being used on own farm (as measured in "ox-days", for example), while *sold* refers to renting out oxen.
- In table 3, inputs from own farm should be included as a category 6 (e.g., corn used to feed the animals). This quantities should still be recorded in section H1, under own use. But, for smaller and less important inputs you may skip them both places (for the sake of consistency and getting accurate income estimates, any own farm inputs must be either included in both sections H1 and I3, or not recorded at all).

J. Other income sources

- Q1: Remittances refers to money transfers from family members that are not part of the household. Thus, if the person earning the money by working far away is still considered part of the household, it is not to be included as remittances but should be included as, for example, wage income, for that person being listed under A1B as a household member (section F).
- Q3: Smaller gifts from friends and relatives should not be included, e.g., a chicken given as part of a reciprocal system. This question is meant to capture the larger unilateral transfers, e.g., a widow that survives on the support she receives from her neighbours (see also section 4.3).
- Q5: Payment for forest services should be specified here. But, the same amount is also asked in A2C in more detail. Make sure that the amounts filled in the quarterly surveys add up to the one reported in A2.

5.8 Attrition (drop out) and temporary absence survey (ATA)

This survey is to be completed for households included in the PEN sample, and that participate in A1/Q1, but miss out on one or more of the following quarterly surveys (Q2, Q3 or Q4). The drop-out may be temporary (one survey only) or permanent (more than one). The ATA survey will be used to see if there are any systematic patterns in the drop outs and absences, and – if so – what adjustments need to be done when inferring from the sample to the full population.

If a household for any reason cannot participate in just **one** of the quarterly surveys, it should still be included for the remaining surveys of the PEN fieldwork. For example, if it misses Q2, it should still be interviewed for Q3 and Q4/A2 (if possible).

If a household misses **two or more** quarterly surveys, it is not part of the PEN sample any more. The income information collected will have too many gaps to be useful in the analysis.

Implementation

The ATA questionnaire should be filled out when a household in your sample cannot participate in a quarterly survey. It should be done while doing the regular quarterly survey (always have a few copies of the ATA survey at hand).

In some cases the household itself can give information on the questions (e.g., if they say they are tired of answering the questions and don't want to participate anymore). In other cases information must be obtained from others: neighbours, relatives, or village officials. Each PEN partner/enumerator should assess who can provide the most reliable information to the questions.

Note that the reasons for dropping out are not necessarily good reasons for not participating in the survey. For example, if they have moved within the village or divorced they should still continue

participating in the survey. As another example, if for Q1 the husband answered the survey, and for Q2 he is not present (out for temporary work or died), you should interview the wife or another adult member of the households. In that case we will not lose the household.

If a household drops out for one survey, every attempt should be made to include them in the remaining surveys. For example, if they cannot participate in Q3, try to get them interviewed for Q4 and A2, if at all possible. (The carrot may be that this is the last time they have to answer all the silly PEN questions!)

Database

Temporary absentees (miss one survey): The information from all the other surveys should be included in the PEN Access database in the usual way, in addition to the ATA survey. PEN will develop methods to estimate (interpolate) the income for the missing quarter, based on the household income from the other three quarters and the seasonal income variation in the village.

Permanent drop-outs: The following surveys should be included in the database: A1 and ATA. There is no need to enter Q1 (and Q2) as this will not be used (incomplete income information). The information contained in A1 will, however, be used to check if there are any systematic patterns in who drops out.

Note: At the time of data entering you may not know if the households have dropped out temporarily or permanently. You may then either enter data from the quarterly survey (just to be on the safe side), or leave it for later when you know if they are permanent drop-outs or not. And, don't delete the survey information already entered before you are sure that the drop-out is permanent.

The bottom line is CQC:

PEN wants to collect high quality data.
Achieving that depends on the three principles
stated in the beginning of the guidelines:

Consistency: *Follow the definitions, codes and guidelines put forward!*

Quality: *Attend to the details at every stage of the research process!*

Completeness: *Complete the complete prototype questionnaire!*

6 General guidelines on fieldwork

This section gives a few general "rules" and recommendations on various aspects of the fieldwork. They are critical for PEN research to achieve a sound research methodology and to get high quality data.

6.1 Selection of study area and random sampling of households

Sampling is normally done in several steps:

1) *Selection of **study area** (district or region).* Most studies will be restricted to one or two region(s) for practical and cost reasons, but including households from a larger area (several regions) is by no means discouraged in PEN, it can in some cases be needed to get sufficient variation in the data and for doing comparative analysis and impact assessments. See section 3.3 for a further discussion.

2) *Selection of **villages***. Normally one cannot include all the villages in the study area, and a selection is necessary. This should be done based on two considerations: include villages that are *representative* for the study area, get *variation across the villages in key variables* to be studied, cf. section 3.3.

3) *Selection of **households** within the chosen villages.* As a general rule this should be done by random sampling, which is critical to be able to derive conclusions from the sample to the total population. There might be exceptions in which one may use stratified sampling. If you think stratified sampling might be a good idea in your study, you should consult your PEN advisor and the PEN coordinator. But, the general rule is random sampling, also because it makes later analysis simpler.

Selection of villages

- **Number of villages:** One key question is if one should go for: (1) relatively few households per village and a large number of villages, or (2) concentrate the households in a few villages. Option (1) is likely to get more variation in the data - and that is often a desirable feature. But it is also the more costly option, and a balance must be found based on available resources and time for fieldwork. The choice also depends on the size of the villages.
- If it is important for the individual PEN study to compare villages, a rule of thumb is that a minimum sample of 25-30 from each village is needed if the village size is in the range of 100 to 500. Further, as a general rule it is rarely to be recommended to include more than 35-40 households from each village. For PEN purposes, however, it is more important to get good variation in the data by including a larger number of villages (option 1) than to be able to compare villages.
- **Distribution of village samples**: Some prefer to have a fixed number of sampled households per village (e.g., 25 for all villages), while others sample a proportion of the households according to the number of households (e.g., 20 % of the households in each village). If one wants to minimize the overall errors of the sample estimates, none of these methods are fully satisfactory, and the optimal sample sizes will be somewhere between the two approaches, and closer to the fixed number for all villages than the proportional sampling.[18]

Selection of households

- The importance of **random selection** of households should be reiterated. Randomness is the *only* way to get a representative sample and thereby draw conclusions about the full population from the sample statistics. Other methods are NOT to be recommended, like a "random walk" through the village (will not get those far away from the village centre), letting the village headman pick the households (selecting the rich ones to give a good impression, or those sympathetic with the government?), calling a meeting and let the attendants be the sample (what about those that are overloaded with work because they are poor?).
- There are at least four ways to generate a random sample of households. Whatever method chosen, keep in mind the *definition of random sampling: each household in the village should have the same probability of being selected.*

1. A reliable **list of households** in the village (e.g., from census or election) is often the best starting point for the **random selection** of households. Number the households, and use a random number generator (e.g., in Excel) to select the households.

[18] The optimal sample also depends on the variation of the variable of interest in the various villages; more heterogeneous villages should have larger sample size. See statistical textbooks for an elaboration.

2. If a reliable village lists are not available, one can plan to interview **every nth household** (n = total # of households/required # of households) as one moves along some dimension of the village. But be careful. If you move along the main road only you will exclude more remote households, and therefore get a biased sample.
3. Using **aerial photos**, and number all the houses in the village, can also be used, provided good photos are readily available.
4. Another more time-consuming, but possibly the most reliable method, is to first make a **village census** and select from that list. A census may also include a few key questions that can be used in the later analysis. See section 6.3.

The concept of random sampling may not be clear to the village head (chief) or others in the village, and notables in the village may feel offended for not being included. Others may suspect that you select household based on who you like or related criteria. Thus it may prove useful to make the selection transparent by having representatives from the village present when you draw your sample. Hopefully those who win the PEN lottery will feel lucky and participate with great enthusiasm!

In some cases it may be important to include the local headman (chief) to ensure his cooperation, and also to facilitate cooperation from those selected. Thus, if the household of the village chief is not among the number of households selected at random, you may still want to include that household in the interviews, and state in public that his household is included. However, that household should not be included in the final analysis. His inclusion in the survey is just to facilitate your work in the village. In other words, you should take advantage of their status, but they should not dictate who should be in the sample.

Households refusing to participate and drop outs

Some of the selected households may refuse to participate for various reasons. Thus when doing the random selection, one should always have a few **extra households** (selected using the same procedures) to replace any households that refuse to participate. One should, however, try to minimize the refusals by approaching them in an appropriate manner and carefully explaining the purpose of the survey. If a large proportion of the households selected refuse to participate this will cause serious selection biases.

A related problem is drop-outs after the survey has started. There are several issues here:

- As suggested in section 3.4, the overall sample size should factor in a 10 % drop out rate. Thus if you think 200 households is an appropriate sample size, start out with 220 households. Once the survey has started, drop outs (in Q2 and afterwards) cannot be replaced.
- The general suggestion is, of course to try to minimize the number of drop-outs as it may cause sample biases. Nevertheless, sometimes drop-outs must be accepted, e.g., the household refuse to participate because some family members just have died, or they move from the village.
- Refer to chapter 5.8 for the specific PEN rules for temporary absentees and permanent drop-outs.

6.2 Rapid rural appraisal (RRA)

RRA - Rapid Rural Appraisal (or Relaxed Rural Appraisal as it also has been dubbed) is an umbrella term for “quick and dirty” methods that can be used to collect the initial information about rural areas (villages). The methods include own observation, talking to village leaders, key informants and ordinary villagers, and arranging meetings for subgroups within the village (focus groups). These interviews and meetings normally have a semi-structured form: the researcher has a few questions and issues he or she would like to raise, but they are asked in a relatively informal way. At the same time, one should not be dictated by these but also follow interesting leads.

While getting into a good dialogue with the villagers and being able to get the useful information requires training, here are some hints that can speed up the process:

- Be prepared! Bring a list of questions you would like to ask (but don't read them out). The PEN village questionnaire provides some tips on useful questions.
- One important issue to be raised is a brief history of the village. What have been the important changes over the last few years?
- Sit down with a group of people and draw a map of the village (if not available): What are the different land uses, who use them (outsiders?), any conflicts over land use?
- Get an overview of the main income generating activities in the village; what is produced, how, where is it sold, etc.
- Making a seasonal calendar of the main activities is valuable information, and also helps to organize your fieldwork. Generating this information is often suitable in a small focus group.
- Talk to different groups of people: the officials, the trader, the NGOs and extension officer, the long-time resident, the newly settled migrant, the rich cattle owner, the poor forest dweller, etc.
- Get out in the field and forest and see and talk to the villagers in their environment.
- Cross check information: particularly during such short visits as RRA involves it's very easy to get misled by just one source of information.
- Make proper notes during the interviews, and discuss the findings in the evening with other members of the research team.

6.3 Village census

A village census is a small survey among all the households in the village. Undertaking a village census is not mandatory in PEN research, but can be a useful instrument for several reasons: (1) it will serve as your sample frame for the random selection of households to be interviewed with the household questionnaire; (2) it will serve as your source of randomized replacement households if some households chosen initially drop out; (3) a census is your means of getting basic information on the entire village for points of information that you might consider vital, not at least for the village surveys; (4) a census will be your means of stratifying the sample selection if you choose to do so; and (5) the census is an ideal way for you to introduce yourself to every household in the community and to explain briefly the goals of the research. This procedure has a high value for encouraging trust and openness among future respondents to produce high quality data.

The desirability of undertaking a census depends on a number of factors. First, to what extent is reliable information about all households in the village readily available and can be used for the random selection of households for the surveys. Second, the geographical extent (scattered or concentrated) and the size of the village, which affect the time and cost of doing a census. Third, the time and resources that are available for fieldwork.

Because a primary purpose of a village census is to be used for the selection of households in the PEN household surveys, it should be done prior to the implementation of these surveys.

As the census if to cover all the households in the village, it has to be very brief, often limited to 10 key questions that can easily be answered. Here is an example of information that can be collected in a village census:

- name of household head,
- sex of household head,
- age of household head,

- level of education of household head,
- caste/ethnicity of household head,
- migrated to village? when?
- total number of household members,
- main source of income of household,
- use of key forest products.

6.4 Designing the household questionnaire

The following suggestions are meant to reflect both the principles which have been used (hopefully successfully) in the PEN prototype questionnaire, and also may be useful for further extending the questionnaire (adding new sections) to fit the particular needs of the individual research project.

Questions to include

- Questions should in general **not require long term, detailed memory**. If questions are too demanding on the memories of the respondents, this will normally reduce the quality of the data significantly. However, the acceptable recall period varies: significant and large events are more easily remembered, e.g., the size of the harvest of the main crop can often be remembered several years back. On the other hand, day-to-day activities such as amount of firewood collected or fruits harvested cannot be remembered over extended periods of time. But we also expect less variation on those activities, so one can multiply firewood consumed during last week by 4 get a good estimate of last months consumption.
- **Open-ended questions** are often too difficult to deal with when analyzing data and end up being abandoned. They are better done in village questionnaires, focus group or key informant interviews. They are particularly useful during the initial stages of research, to learn about the area and prepare a systematic questionnaire. A test question to ask when deciding on whether to include new questions is this: can the answers be categorized and used in statistical analysis? If not, leave it out.
- **Avoid qualitative questions**. They are difficult to use in quantitative analysis – which is what the data are to be used for. They are better suited for a few in-depth interviews. If included in the household survey, think through carefully and pretest the wording, and use a few predefined categorical responses.
- Do you really need **consumption expenditure data?** If you are to check that *consumption + net savings = income*, then you must carefully record all consumption items, not just a few (then it's useless for that check). Moreover, the recall period should be short for consumption expenditures (a few weeks, and not years). Having said that, some simple questions about consumption can be useful to check the consistency and reliability of income data. If you suspect some income sources to be forgotten or underreported, some simple questions can reveal this: "How often do you go to town to make purchases? How much on average do you have to spend?" Some simple calculations can then reveal any underreporting of cash income, and follow up questions should be asked.
- **Family labour data – are they needed?** Labour data are useful for a number of purposes, e.g., calculating the return to labour in different activities the household engage in. But, it takes a lot of time to get reliable data (ask for each household member, for a number of tasks, and have short recall periods). We have therefore chosen not to include it in the PEN prototype

questionnaire. Also remember that labour data are not needed to calculate net income (income is the value added of labour and capital).

- Each question takes **time**, and time is costly – also in reducing the quality of the responses to the other questions. Thus, for each question included the test question is: how am I going to use this information? If you don't know precisely, drop the question!

Formulating the questions

- The questions should aim to gather disaggregated or decomposed data, which can then be aggregated by the researcher. *Disaggregation is critical to get reliable information.* For example, a respondent is unlikely to know what the total income from agriculture was in the just ended season. But if asked questions about the quantities of each crop harvested and the price at which they were sold, quite accurate responses can be obtained. More generally, income can be decomposed in a number of ways, which has been done in the PEN prototype questionnaire:
    a. Crop, product or type of work
    b. Plot, or type of forest
    c. Household member (e.g., wage income)
    d. Season
- Design your questionnaire in such a way that respondents can report measures in **locally understood units** (e.g., in scorch carts, buckets, bags) rather than expecting them to report in unfamiliar metric units. A small units survey can then be conducted later/concurrently to establish conversion factors for the local units into standard measures. The units used in the data entry should be reported to the PEN coordinator and assigned a code in the PEN code list.
- When structuring the questionnaire, **pre-code** as many responses as possible but leave room for other responses. The codes should be in clear, should avoid overlap among the choices, and should include all relevant alternatives. Check and refine the codes during pre-testing.
- Be **concrete**, and avoid hypothetical questions. For example: don't ask: "how will you respond to a food shortage". Instead, ask if they experienced a food shortage in the last 5 years and what they did in response to the shortage.

The questionnaire design process – pre-testing

- Drafting of the questionnaire can be done before arriving on site, but input from concurrent PRA/RRA exercises, key informant interviews, reconnaissance visits and pre-testing is critical for formulating relevant and appropriately phrased and unambiguous questions and also for formulating the codes for likely responses. The questionnaire can also be refined with the help of people on site or other researchers who have worked in the same area.
- **Test the questionnaire** with at least 6-7 households in the area to make sure all the questions work. As you test the questionnaires, record the time it takes to answer the whole questionnaire. A general rule the pre-testing should never be done in the households included in the survey, and possibly not in the same villages either (although this might be impractical in some cases).
- Work thoroughly with the enumerators on **translating** the questions into the local language, or into a locally well-understood language until all questions can be put across clearly and correctly. Use two enumerators to do double translations. One translates from English (or any other international language) to local language, and the other (independently) re-translates

from local language back to English. If you recognize your question, the translation is acceptable.

- You, as a PEN partner and field research supervisor, and each of the enumerators should test the questionnaire until it is thoroughly revised to your complete satisfaction, and until each enumerator thoroughly understands every question. You should attend at least one pretest interview with each enumerator to be sure that the interview technique is up to your standards. If not, you should replace the enumerator. Remember: The cost of finding out that an enumerator is incompetent can be unacceptably high if the discovery has been made after the enumerator has interviewed (say) 50 households. Not only are the data (possibly) lost, but you may be required to interview different households to compensate for the lost data.
- Plan for at least one week of pretesting and translating before you plan to start the actual data collection.

Some practical hints

- Questions are better drafted into **tables** for ease of recording, coding and data entry.
- **Spiral binding** of questionnaires helps in getting organized and avoiding loss of some questionnaires or questionnaire pages. It also provides easy access for cross checking during data cleaning and entering. With several household and village surveys things easily get messy. Thus you may have one folder for each household, containing all the questionnaires.
- **Colour code** the different quarterly surveys to avoid confusion.

6.5 Implementing the surveys

Approaching the village and the households

- When approaching the community the researcher should be as **truthful** as possible about the purpose of the research and survey, and avoid raising unrealistic expectations. It should be made clear when approaching communities that the exercise is for purposes of research only.
- As a general rule, you should ask for **permission from local authorities** to conduct the questionnaire. When you reach a new household, it can then sometimes help to explain that you talked with the authorities and they agree on the research being conducted. But also remember that making reference to local authorities may diminish your chances of getting reliable answers, for instance, you may be mistaken for being a tax collector! In many contexts it may be appropriate to organize an open village meeting to explain the purpose of the research. But, customs vary, and in some places a written permission may not be needed and take too long time.
- If possible, plan the survey so that it coincides with the **off-peak period** when households are not too busy. Attempting to do a survey when people are busy may run into problems as respondents are hardly found at home, and even when they are found they may not have time to adequately respond to your questions.
- **Booking of time** can often be useful, e.g., see the households that you will interview the day before and make sure they are available. If they are busy, ask for an appointment in a more convenient time, and come back as people request.
- During the first visit, tell the households you will be **coming back** for more surveys.
- When making an appointment with the household and/or when starting the interview, the following points should be communicated with the respondent:

- The interview will be with the head of household and the spouse (assuming there is a spouse).
- Only members of the household should be present during the interview.
- The purpose of the research, the activity is scientific and is not a development project, etc.
- Confidentiality will be assured.
- The length of the interview.
- Be also prepared to explain in simple terms why some households have been included and others not in the survey. (Some respondents will then feel lucky!!)

During the interview

- Show a **genuine interest** in the respondents. This is both good manners, and will increase the reliability of the answers.
- Some **small talk** may help to get the questionnaire going; too much of it will extend the interview period excessively. Use your intuition to strike a balance! It is important for the people being interviewed to feel comfortable about the situation.
- Keep the **interest and attention** during the interview. Sometimes a small break may help. If the respondent gets bored his/her main strategy easily becomes one of giving some quick but *apparently* trustworthy figures to quickly get you out of the hut!
- It is crucial that the questions are not formulated as **leading questions**. For example, if the respondent cannot come up with answers, there is a deeply ingrained tendency for untrained enumerators to suggest possible answers. This is a fatal error unless it is eliminated before the research ever starts.
- Make sure there is **no prompting of possible answers** by non-enumerators and non-respondents, e.g., forestry personnel that often accompany researchers to the field, or visiting neighbours that just would like to assist. Thus during the interview only members of the household and of the research team should be present. (There may be some situations where government protocol requires the accompaniment of an extension agent in household interviews. Do your very best to discourage the agent from attending, pointing out that the presence of government officials is known to distort the answers of respondents, and scientific procedure requires that respondents be free of any constraints or self-consciousness. If this does not work, then ask that the agent be completely silent during the entire interview.)
- When using recall questions, try to **link this to particular events** in the family, village or country. Example: some questions in the PEN prototype questionnaire ask about the forest resource base 5 years ago. If the family has a 5 years old daughter, one can then ask about how the situation was when she was born compared with the current situation.
- Some households may tend to provide answers that reflect village averages, rather than their own household (e.g., they feel they are not representative). Get the **household specific information!** Make sure that it is clearly understood that you are interested in data for this household, even though it's not an average one. (It's the non-average households that give you statistical significance in later analysis!)
- All the respondents should be **asked the questions in the same manner**. Thus insist that the enumerators (and yourself) are following the wording of the questionnaire. Shortcuts are easily done when the interviewer has become familiar with the questionnaire. Resist the temptation, and ask respondent no. 247 the question in the same way as for no. 3.
- Work on **interviewing skills** (of yourself or the enumerators). The interview should be more like dialogue and not an interrogation. Try to use neutral yet non-mechanical questioning.

Explain to enumerators the importance of not prompting answers, i.e., ask in such a way that the respondent does not get the impression that a particular answer is expected or preferred.

- **Make it flow:** The quarterly questionnaire is more like a data worksheet than questions that are to be asked directly. This is typical for many questionnaires applied in the field, and this put some more responsibility on the enumerators to ask them in a nice way. For example, the section B in the quarterly questionnaire, the questions could be asked in about the following way:
  - Did any members of your household collect any mushrooms from the forest for own use or sale over the past month? If yes:
  - Who collected them? (column 2)
  - Where did you/he/she collected it (column 3) (You may also ask about the ownership – column 4), or you know from other households interviews the ownership and fill that out)
  - Was it for your own use or did you sell it?
  - How much of was for own use? (column 5)
  - How much of did you sell? (column 6) Etc. etc.
- **Check consistency of responses:** Some responses may, when combined, not internally make sense. Check this during the interview, or when going through the questionnaire in the evening.
  Example 1: A respondent reports large veterinary expenses and significant pasture areas, that may go ill along with the fact that (s)he only reported to hold five heads of cattle. It will thus be relevant to go back and question the earlier response on livestock holding and resolve the paradox.
  Example 2: By the end of the interview it will appear as if the respondent households were earning basically no monetary income. That may seem unrealistic, in particular vis-a-vis the perception of degree of wealth you have from seeing the house where the respondent lives. Try the following check. Q: "How often do you go to the local market to buy household necessities?" A: "Once a week" Q: "So, how much do you then usually have to spend?" A: "Oh, now things have become very expensive! On average, 15 US$". Q: "Hmm, that means that you have to spend at least 60 US$ per month on basic necessities. With the limited income sources you have just mentioned to me in the interview, how can you actually afford that?" Simple add-on checks of this type can sometimes resolve paradox inconsistencies due to misunderstandings and (accidental or deliberate) omissions during the interview.
- **Emphasize key questions:** An aim of any survey is to minimize the overall errors, e.g., get a good overall income estimate. This means you should focus on the important activities, and not be spending less time on the minor details. For example, from your previous village and household interviews, you know that cattle ranching tends to be the key income-generating activity. You may then have to invest more specific time to be sure to get things right on cattle holding.
  Example 1: Q: "How many heads of cattle do you have?" "A: Eeh, five". Q: "Hmm, that is not a lot, considering the 20 ha you have in pasture" A: "Yeah, I'd like to buy more, but cattle is so expensive, and we cannot get any credit from the government out here!" Q: "So, those five heads, is that adult ones only, or also calves?" A: "Adults only; I also have three calves" Q: "Is that from this adjacent plot only, or also from the remote plot your wife is managing?" A: "Well, from this one only: actually on the remote one we have another ten heads; but we use them only for milk production."
  Example 2: In some areas remittances are absolutely vital as monetary income sources, but often respondents are reluctant to tell about them straight away to a stranger, given the amounts of money being involved: Q: "Did you receive any remittances over the last three months" A: "Remittances? Me? No." Q: "Hmm, but you said before that your one son is working in city/country x. Doesn't he want to help his poor parents sometimes? So, he is not a very good son, eeh? ;-)" A: "Well, he has a lot of costs of his own - life is expensive in the US!

But actually two months ago he sent us US$2000 because the truck broke down/ because my wife got ill, and we could not cope with the situation on our own." In other words, don't be afraid to "go fishing" if you think there might be more under the surface than appears at the first response. It does not always pay off, but sometimes it is amazing what hidden things it can bring to the light of day!

- **The bottom line is this:** Don't be naïve and use your insights and sound judgements. Don't just mechanically record answers assuming the respondent is telling you "the truth, the whole truth, and nothing but the truth", as if he or she had zero vested interests in biasing any of the answers. Try to understand the livelihood logic of the household responding to you, and try to understand their self-interests. If some things appear unlikely or inconsistent, raise the issue in a polite but direct way during the interview, check with other sources, and go back if necessary. And, teach enumerators a similar approach.

Checking, coding and data entry

The job is not done when you or the enumerator return from the household with the questionnaire filled out. There are several important steps along the road to a complete and quality-controlled data set ready for analysis. Each questionnaire therefore has a first section labelled "control information", which describes the four main steps long this road:

1. **Checking questionnaire:** check each completed questionnaire as soon as possible and ensure that enumerators are doing the right thing, ask them to correct any obvious mistakes in consultation with the researcher. The following procedure has proven useful: all forms are checked for completeness and accuracy in the afternoon or evening by the enumerator. In this way errors are discovered while memories are still fresh, making it much easier to get the correct information the next day if needed. Then, once the enumerator is satisfied, (s)he signs on the front page of the questionnaire (in the "By who?" column). The form is then handed over to the field supervisor (in most cases: the PEN partner) for recheck and approval. This helps encourage responsibility and screen out dishonest enumerators who write fiction.
2. **Coding questionnaire:** The coding may be done as part of the checking of the questionnaire (1), as a separate exercise, or as part of the data entry process (3). It partly depends on how the fieldwork is organized, e.g., if you have a field supervisor (s)he may have the responsibility for the coding.
3. **Entering data:** Data entry is done using the PEN database entry form, designed in MS Access (see separate user guide). Again, the responsibility and organization of the data entry will vary from study to study. In some cases the PEN partner may do it, in other cases someone may be hired for just this task.
4. **Checking and approving data entry:** It is important to stress that in the end the PEN partner is responsible for approving the data set. The checking should look for gaps in the data, check inconsistencies, and identify outliers or extreme values (which *may* be due to either typos or incorrect reporting in the questionnaire).

We would like to stress a couple of general points:

- Do the above steps **as soon as possible** for a number of reasons. First, common mistakes, misinterpretations or omissions can be identified and corrected immediately. Second, one may revisit households if necessary to fill in gaps. Third, one may identify enumerators not performing well (or even cheating), and take appropriate actions. Fourth, one may in fact do some preliminary analysis of the data (say of the first 1-2 rounds of the survey), and develop new hypotheses, which may be checked in, for example, new focus group meetings.

- The PEN partner should be actively **involved** in all stages to maintain the quality standards: provide guidance and correct errors immediately, and reduce the temptations of making short-cuts.
- Take the necessary measures to keep the household specific information **stored safely.** The information can, in the wrong hands, make respondents a target for theft. Thus the questionnaires should be locked-up, and the information in the computer stored safely. In one sensitive PEN area the first page with the household specific information was detached from and stored in a bank.
- Also keep back-ups of the questionnaires. If the information is not entered into the computer soon after being recorded, you may consider making **back-up** copies of the questionnaires. And, always make regular (preferably daily) back-ups of the data entered. PCs do crash, and get stolen! There are several stories of entire surveys being lost – and there are very, very few good excuses if that happen.

Some practical hints

- A **small gift** for households participating in the survey will always boost moral, especially in the case of repeat surveys, and help avoiding respondent fatigue and withdrawal from the survey. Although in some cultures cash payment may be preferred, as a general rule the payment might be in kind rather than in cash. It should be of practical use and not be overdone but reflect the time spent on the interview. For example, a kg of sugar and a bar of soap can be an appropriate present – unless the household produces sugar itself and seldom uses your type of soap! You can also think of a community gift (notebooks for school children, a soccer ball, etc.). Or, in some cases buying some products (chicken, basket, etc.) from the household may also be appropriate and is a sign of appreciation of their work (and don't be too tough in the bargaining!). Bottom line: what is considered appropriate varies a lot, so use your sound judgement. Beware that gifts create expectations – both for the next survey round and for future researchers, so use them with care.
- **Getting involved** in a few community activities always results in better relations with respondents and help to reveal the more sensitive facts and uncover the real situation in the village. Go with them to church/mosque one day, attend a local funeral/wedding, play soccer with the school children/youth, play chess with the old men, anything!
- Rather than leaving much of the fieldwork to enumerators, it is in the interest of the researcher to spend as much **time** as possible in the **field** during the surveys, to increase interaction with enumerators and respondents. Participants always take a survey more seriously if the people behind it show interest through continuous presence. Concurrent activities such as PRA/RRA activities and other exercises to collect village level data always help to give weight to the household survey.
- Underline repeatedly the strict **confidentiality** of household-level information. As part of this, you may want to keep a certain **distance from government** officials, e.g., they should never participate in PEN household or village interviews! You may also think twice about staying in the house of a government official or a government guesthouse. The PEN partner and members of the research team should in general avoid to be seen as a prolonged arm of government.

6.6 Enumerators

Normally, you are not doing all the questionnaires by yourself, so having good enumerators is a key success factor. In fact, the experience from the first PEN surveys that have started

demonstrates that getting a good research team together is possibly the single most important factor for the success and quality of your research!

To successfully implement the PEN prototype questionnaire requires a solid understanding by the enumerators on the purpose of PEN, for example, to ask appropriate follow-up questions to get the information intended.

- Careful **recruitment** of enumerators is vital. Set a minimum requirement, and get applicants from or familiar with the study area. Conduct interviews with applicants, be as transparent as possible – employment issues are usually sensitive in rural areas if recruiting locally.
- While knowledge of **local language and culture** is a prerequisite, using enumerators that the respondents know directly (or know their family) is often a disadvantage: respondents may be less willing to reveal sensitive information to them. But this varies: in some places villagers are **honoured** to be visited by outsider, while in other places they get **suspicious**.
- Intensive **training** and follow up of enumerators always pays off, and is a vital step in familiarizing and pre-testing the questionnaire. A simple way to start this training is for the enumerators to apply the survey among themselves through role play, prior to the pretest.
- Enumerators will usually require a lot of **practice** before they can get this right; let them practise on each other in the relevant language, then with households that are not in the actual sample.
- During the course of the survey, any **new enumerators** that are engaged to replace some who might have dropped out must be sufficiently trained before they can take over.
- It is very important to keep the wording of the questions constant. Insist in that all enumerators use the **same wording** for the questions
- All enumerators should go through **a week long training programme**, depending on their previous experience (up to two weeks if inexperienced). A training programme should include the following agenda items:
    1. PEN and your research project (use material found on the PEN Web site):
        a. Purpose and objectives
        b. Definitions, including income, forest, …
    2. Go carefully through the guidelines, particularly this section
        a. Approaching villages and households
        b. The interview situation
        c. The importance of accuracy
        d. The method of disaggregating the question
    3. Define and agree on key terms in local language (household, forest, income, …)
    4. Role play of questionnaire
    5. Test the questionnaire and the enumerators in the field

6.7 Valuation and pricing

It is not uncommon, especially in more remote areas, that the majority of goods produced and collected by a household are destined for consumption directly by the household (auto-consumption, or subsistence consumption). Thus, there is no explicit market price for valuing the consumed quantities of goods, but still we need to assign an implicit price in order to calculate the value and compare "apples and oranges". The selection of pricing method can make a large

difference in terms of our overall income results. It is thus a key issue to think about carefully, requiring some economic reasoning and a lot of common sense. **This is an issue you always need to consult with your PEN advisor and/or coordinator!** The applicable methods are, in rough order of priority:

1. **Farm-gate (local) prices**: As a default rule, one should use farm/forest gate or local market prices (i.e., within village), if available. Exceptions are if these markets are extremely "thin" – they are only traded exceptionally compared to a large share being consumed (say, less than 10% of the total), so that the available price gives a biased picture of local value.

2. **Barter values**: The barter value is the implicit value from trade with a *market* commodity.
   Example: If one kg of mushrooms is usually exchanged between households for 2.5 kg rice, which has a market price of Shilling 20, then the monetary value of mushrooms is Shilling 50 per kg.

3. **Substitutes**: When trade (either in a market or through barter) is not common, one can find the implicit value by considering close substitutes that has a local market price.
   Example: Dung is being used for fuel together with firewood, and there is a market for firewood but not for dung. The question is then (asked a few households or in a focus group meeting): what would be the costs of buying firewood that would yield the same amount of energy as one unit of dung?

4. **Willingness to pay:** Asking respondents about their willingness to pay is a common method in environmental economics under the heading of "contingent valuation" (CV). Doing a proper CV would take too much time, and a *collective CV* is therefore suggested: Get a group of villagers together and ask them about their "willingness to pay" for the particular product. One can use either monetary units, or the equivalent in the staple food.
   Example: There is no market for firewood, and rice is the staple food. The villagers are then asked: *How much rice will you give me for one headload of firewood?* Ask them to reach a consensus price.

5. **Distant markets prices:** Prices at more distant market deducted for transport costs can also be used. One problem with this approach is that the price net of transport costs might become negative. That might in fact be the reason for the non-existence of a local market!
   Example: There is no local market for charcoal in the village, but a bag is sold at 200 shillings in the district capital. If the transport cost is 50 shillings, the price to be used in the survey is 150.

6. **Value of time:** The time used multiplied by local daily wage rate (serving as the shadow value of labour) can be a measure of the value of the product. But this approach also raises some problems: (i) exact measuring of the time spent (including problems of "multi-tasking", e.g., collecting firewood on the way home from an agricultural field), and (ii) determining the opportunity cost of labour (who does the forest product collection in the household, peak-slack season, opportunities in the labour market, etc.). (iii) It is implicitly assumed that the resource as such fetches no "rent", i.e., it has no value beyond the labour used to collect it, which will only hold for marginal or inferior products.
   Example: If the husband goes for two days to collect 10 kg of mushrooms, and alternatively could have sold his labour in the casual labour market for 15 shillings a day, the minimum price per kg mushroom should be set to 3 shillings: (2*15)/10.

One should note that the same product may be sold at different prices, especially depending on season. Both agricultural and forest products tend to be subject to accentuated fluctuations in prices. This will be captured by doing the quarterly income surveys in PEN, and it is important to ask explicitly about prices in each survey (and not use the ones from the previous quarter).

Another issue is when the products are sold in different markets, e.g., firewood sold dearly to traders but cheaply to fellow villagers. One should in those cases report an average price.

6.8 Security and safety

The safety of yourself and your co-workers should have no. 1 priority. Fieldworkers are often in a vulnerable position, being strangers in the area and often considered "rich" and therefore a valuable target of crime. The situation varies considerably across countries, so the rules of thumb may not apply equally to all the locations. But everyone should take the necessary precautions to minimize the risk by following some basic rules:

General

- Make careful choices about study areas, avoiding/abandoning if possible those that are conflict-ridden or where conflict could emerge.
- Give full details of medical aid, next of kin, contact numbers, ID number, etc. to the local partner institution, or someone that you are working with (but do not accompany you to the field). They should also get the detailed plans of your field trip (where and when).
- Keep the local headman, village chief, police station or other relevant authorities informed about you stay and movements.
- Fieldwork should not be undertaken alone. There must be two people present, preferably three, of whom at least one should be male.
- Bring a cell phone, if the area is connected.
- Take into account that "Murphy's Law was written in the tropics": many plans in study sites can go wrong, so make some contingency arrangements whenever possible. Thus go through with your research partners some basic safety routines: *What do we do if: the car breaks down, a research team member is attacked or get sick, etc.*

Crime

- If possible, ally with some local family in each village you work so that you locally "belong" under someone's custody.
- Make some security assessment when choosing in particular accommodation sites (a local trustworthy family may be the best option).
- Take seriously any threats from individuals you might receive at your study sites.
- Avoid as much as possible openly exhibiting valuables, which might label you as a target for economic crimes.
- Beware of participating in "dubious" social gatherings (e.g., with lots of alcohol involved) after dark when alone and unprotected.
- Be careful when hitchhiking, or offering lifts to strangers, especially when alone, in remote areas, and/ or after dark. In many areas the rule should be to never give lifts to anyone not associated with the project.
- When in transit do not stop your vehicle in remote places. If the road is obstructed by rocks, tyres, poles or holes, reverse immediately, or drive through. Do not get out to remove the obstruction, this may be a trap set up by criminals.
- Be careful driving after dark. Wrap up your fieldwork early enough to reach home before dusk (this may mean leaving out that last questionnaire, transect, etc.).

Medical

- Driving is probably the activity that entails the highest risk. If you rent a vehicle (and driver), you set the rules. Tell the driver to slow down, or set maximum speed limits.
- First Aid kits must be taken on all field trips.
- Carry ID and medical aid details.
- Make contacts with local doctors or hospitals to check out the assistance available in cases of emergency. Make contingency plans.
- If someone gets sick, another person should take charge. Don't leave it to the sick person to decide, often they are unable to make rational decisions.