4. DATA COLLECTION METHODOLOGY

As noted in section 2, the specific objective of this research was to assess the impact of seed systems on farmer welfare and on agricultural biodiversity. Towards this objective, once the survey site and collaborators were identified, the next step was to develop and implement a methodology for the collection of pertinent information. In this section, the manner in which this was done is described.

Note that given the desired focus of the study, the project team felt that a mixed quantitative and qualitative approach to collecting data was appropriate and necessary to meet the objectives of the research. To this purpose a series of surveys[14] and focus group sessions at household, community and market level of analysis were agreed upon and will be further described in section 4.4.

4.1 The data collection strategy

A key to the success of the research program was to ensure that the data necessary to answer the research question was collected in a proper manner. Correspondingly, it was crucial that the project team carefully defined the principal set of variables needed for the analysis and the mechanism by which these would be measured. Three sets of variables of particular importance given the objectives of the program are those related to measuring the seed system, agricultural diversity and farmer welfare. Along with these, it was essential to collect other variables considered to influence these key outcomes. We proceed by first discussing the information required and then the strategy to obtain that information.

4.1.1 1nformation requirements

To get a measure of the seed system, for both supply and demand side, the team agreed that the following pieces of information were needed:

-sources and availability of improved varieties (e.g. dissemination of extension packages or improved varieties at regional level);
- index of market penetration (e.g. quantity, quality and prices of seeds and grain available);
- availability of and accessibility to diverse varieties in informal system (through exchange or in seed markets);
- transaction costs associated in getting a specific variety (existence and use of social capital in getting seeds and information, existence of formal negotiation agreements, usual source of purchase, source of information, costs and means of traveling and of transportation, distance to cover, difficulty in obtaining seeds of desired varieties; costs and availability of other inputs, value and availability of labor forces etc).

On the demand side of the seed system, it was important to take into account that the farmers' utilization of CGRs is driven by production conditions as well as socio-economic conditions besides issues of availability and accessibility also linked to the supply side. Concerning farmers' preferences and constraints, particular regard was devoted to data needed for quantifying the level of farmers' welfare. The neecessary information to run the relative analysis was considered to be the following:

[14] Details on surveys characteristics, and design are given in next section

- measure of type and amount of land ownership, use and rental;
- employment and income generation;
- measure of household assets and tenure;
- crop production by plot and as well as net return to crop;
- vulnerability, crises and copying strategies;
- off-farm income, finance and access to credit etc.;

Other data, at a different scale of analysis, was necessary to collect in order to control for effects of environmental and agro-ecological conditions, presence of risk and management strategies, presence of infrastructures, external interventions in the area, access to credit, presence and measure of social capital common to all household resident in the same village.

Furthermore, considering that market conditions are likely to have an important influence on farmers' seed selection decisions and thus on the HCS project impact, we also needed to collect data for prices and quantities of sorghum and wheat varieties sold in market outlets relevant and accessible to surveyed HHs. Whereby market outlets we mean, local institutions consisting of a conglomeration of buyers and sellers meeting on a periodic basis to exchange agricultural commodities and seeds.

Last but not least, data related to crop and varietal diversity were also collected in order to be able to measuring on farm crop biological diversity. These data were used to construct three indices of diversity, all adapted from the ecological literature[15]. The first index considered, was a count of the total number of crops (or varieties) that the household reports planting over the season of interest and it expresses the *richness*. The second index considered was the Shannon index which expresses proportional *abundance or evenness*, accounting for the land shares allocated to each crop as well as the number of crops (or varieties). The index gives less weight to rare species than common ones, but is more sensitive to differences to small degrees of relative abundances than the Simpson index, another widely used evenness index measure of diversity (Baumgartner, 2004; Magurran, 1988). Finally, the Berger-Parker index of inverse dominance reflects the relative *abundance* of the most common species (Baumgartner, 2004; Magurran, 1988), or the most widely grown on each household farm. To calculate these indices we needed data on numbers, names and area in each crop and variety, taking into consideration the possibility of mixed cropping and intercropping.

Moreover, as these data would have been complemented by the agro-morphological characteristics, information on variety traits and characteristics was also considered essential.

A critical issue to consider in collecting household data is the appropriate unit of analysis to collect data by. Two considerations are important: 1) the unit at which farmers will be most familiar with and thus be most accurate in responding to and 2) the unit required for the type of analysis planned from the data. These are not necessarily the same and careful thought needs to be given about which questions should be asked for individual plots or crops (or crops by plots), the entire farm operation, or household level. A summary of the information that was collected in the household survey, together with the unit at which the data was collected and the analytical purpose for which was being collected can be found in annex 2.

4.2 Components of the data collection strategy

The strategy adopted to collect information described in previous section was that of using a series of five different but linked surveys:

1. Household (HH) survey
2. Community survey

[15] Details on the formulae for the indices are given in the annex 1

3. Market survey
4. Agro-Morphological survey
5. Focus groups

4.2.1 The household survey

The household survey was intended to be the main instrument to collect the required data, which would then be supplemented and complemented by the other surveys and studies. Taking into account the length and number of questions needed to collect necessary data, it was considered appropriate to split it into two different rounds of data collection. The first round to be done after the Meher (main season) planting and the second to be done after harvest of the same crop. This way planting and harvesting decisions taken over the entire production year would be captured.

Several discussions were held on the optimal time to collect the various types of data. Certainly the survey needed to be conducted after planting as it would focus on the actual crops and varieties selected for the production season. In the Hararghe area, this meant after planting for both the Belg and Meher was completed which would occur after July. At the same time considerations of peak labor requirements on farm were important, as farmers would be unwilling to take the time for a survey during busy periods. Additionally a consideration was the timing of the availability of the survey team, which was drawn from the Alemaya University community, who would not be available for such work while classes were in session.

4.2.2 Community and Market Surveys

The community survey was implemented in the same communities in which the HHs survey respondents resided, and collected information on the presence of infrastructure, services, distance to main markets, land use patterns, seed system interventions and general information felt to be potentially important determinants of the functioning of the seed system.

The market survey was designed to provide complementary information gathered on the seed supply in terms of prices and quantity sold in the markets distinguished by varieties and by quality (i.e. seeds vs. grain) for sorghum and wheat. At the design phase of this survey, questions on price differentials between grain and seed were addressed.

4.2.3 Agro-morphological and Focus Group Surveys

The agro-morphological survey was intended to validate CGR information gathered from the HH survey through creating a link to variety names, identification and validation of traits. Since replicated trials were not possible within the framework of 2002 work it was considered only possible to obtain qualitative estimates and therefore data collection would focus on traits of high heritability which show little Genotype by Environment (*GxE*) interaction. Quantitative traits such as number of farmers' varieties, inflorescence length, width and plant height as well as qualitative and semi-qualitative characters (i.e. those that can be expressed in all environments and can be observed such as seedling vigor). Synchrony of flowering, Midrib color, plant color, nodal tillers, Panicle type, awndness, glume color, glume cover, Kernel

color, Endosperm texture, Presence or absence of sub-coat, thresh ability and others) were the objects of data collection strategy. These estimates were expected to provide important information on the amount of diversity in the study area *vis-a-vis* Ethiopia as a whole (comparison with IBCR data) and on any inequalities within the study area (diversity or traits associated with specific PAs or Woredas). Only varieties where focus group results suggested that names were used consistently were considered to be validly subjected to this analysis.

Moreover the agro-morphological characterization was meant to analyze the effects of HCS activities on infra specific diversity of sorghum and wheat. The objective was to investigate what effect HCS activities had had on diversity within the varieties on which they have worked. A further question to be investigated was whether HCS activities had changed the numbers or characteristics of other local varieties of sorghum and wheat grown in their areas of operation. These changes might affect only households to which HCS has distributed seed or other households in the PAs where HCS has worked.

The focus groups were important to validate estimates of the amount and distribution of diversity in sorghum and wheat in the study area and the extent to which these differ in different PAs and woredas in Hararghe. The validation of farmer names and confirmation that different farmers use the same names for the same varieties in different villages, PAs or woredas is a necessary step in a comprehensive seed system analysis. Studies have shown that in some cases farmers may use the same name for agro-morphologically different varieties. In others, they use different names for agro-morphologically identical varieties (Jarvis *et al.*, 2000).

A number of participatory procedures have been shown to provide the information necessary to validate farmer variety names (King, 2000). In this particular case, it was agreed to use focus group discussions which included semi-structured questionnaire and open discussions to confirm (or not) variety identities, differences and similarities in names and qualities of the varieties, and key characteristics of specific named varieties. Training in focus group participatory procedures to obtain the information required would have been provided by IPGRI using experts from Nepal[16].

Focus groups serve to elicit and compile information on how farmers characterize, distinguish, and value land races (e.g. use of four-square method, diversity blocks, gender disaggregated assessment of consistency in farmer unit of diversity or farmer named cultivars, social seed exchange and distribution analysis). More specifically, participatory appraisal and semi-structured interviews, based on open frameworks and two way communication to adapt questions to responses, would help assessing local taxonomies to determine farmer definition of varieties, crop properties, and crop values. It would also allow the identification of farmer strategies, and decision-making which shapes the management of crop diversity, provide insight into group decision-making, priorities, and values, generate information about local environment and social systems, perceptions of responsibility, ownership, physical or social boundaries, environmental factors and agricultural activities. The exercises were to be complemented by transect walks to delimit the main agro-ecological zones within a community, distinct in terms of ecological, agricultural and/or social features.

4.3 Survey sample design

[16] Analysis procedures are summarized in A King (o*p. cit*.) and Rana *et al.,* (2000).

The design of the household survey sample forms the basis for the other components of the data collection strategy. The community surveys and community focus groups are conducted in the same communities included in the household sample. The sites for the market survey were based on proximity to the surveys households and the agro-morphological data was collected for a subset of surveyed households. For this reason, the household sample design was particularly important and is next discussed in detail.

4.3.1 Household survey

Our sampling strategy was built around two main principles: minimizing sources of variation not related to seed systems (e.g. agro-ecological zones) and stratifying the sample by proximity to market and participation in the HCS seed program, which are expected to be major determinants of variation within seed systems. The reason for this is that we wanted to survey households similar in every feature except whether or not they had participated in the HCS seed system intervention. The experience of a lowland HCS farmer is not comparable to that of a highland non HCS farmer. Our sample design includes considerations of criteria at the woreda (county), peasant association (municipality) and household level.

A first step in drawing the sample was to get very clear and detailed information on how households and communities were selected for participation in the HCS project and a sense of how these were distributed across agro-ecological characteristics such as highland/lowland, degree of rainfall, length of growing season, soil and topography. The HCS project areas are located in three major agro-ecological regions – lowlands, midlands and uplands. A major difference between them is their elevation which determines, in turn, also their rainfall and cropping patterns. Different crops and varieties are used in different zones and their planting and harvest dates vary as well. The sample was limited to three woredas- Chiro, Meta and Dire Dawa. The intention was to identify areas within these three Woredas that had similar agro-ecological conditions. Secondary data on agro-ecological conditions was not available at the necessary scale of analysis however, (e.g. Peasant Association) consequently the sample ended up with more variation in agro-ecological conditions than desired. Although ostensibly only PAs considered highland or midland production zones were included, considerable agro-ecological variation existed even within this range. This was controlled to some extent by the collection of household and PA level information on agro-ecological conditions which was used in statistical analysis, however better control of this factor in the sampling stage would have been preferable. A total of 15 peasant associations (PAs) in the three woredas were selected using the criteria of highland to midland production zone.

Once the HCS-project PAs were selected, the next step was the identification of 15 comparable non project PAs within which to select households operating under similar agro-ecological conditions, in order to have control groups for measuring the impact of HCS project. Comparable non-project PAs had to have sorghum and wheat as primary crops and problems of seed insecurity. To this purpose we gathered information, from HCS, regarding agro-ecology, market access, population density, extension program and other NGOs activities, cropping patterns and seed insecurity issues for possible comparable PAs. With this information available 15 comparable non project PAs were selected.

Having identified woredas and PAs in which to select households (HHs), next crucial step was the selection of the households to interview. The total number of households that had participated in the HCS seed program up until, and including 2001 was used as the sampling frame for project participants. The sample drawn represented approximately 5% of the households that participated in the HCS program in 2001 in the selected communities (360 households of a total of 7257). In addition to HCS project participants, we wanted two types of control groups: those who did not participate with HCS but lived within communities where the program was implemented, and those who did not participate and live in the 15 comparable communities where no program was implemented. The first sub-group (*from now on:* non-HCS1) would have provided a source of control for spillover effects and

thus for measuring the degree to which project impacts have been diffused to non participants, as opposed to the real control group (non participant in non HCS-PAs, *from now on:* non-HCS2) which should be uncontaminated by project effects. This of course was assuming that spillover effects would occur only within and not among communities.

Considering that the study timeframe and budget availability dictated the total number of households feasible for sampling with a total of 720 households, the remainder of the sample was divided between the two types of non participant groups. The final composition of the household sample consisted of 360 HCS project participant households, 180 non-project participant households in PAs where HCS had program activities and 180 non-project participant households in non-project PAs. For each of the project PAs, 24 project participants (360/15) were sampled, while for each project PA, 12 non-participants (180/15) were sampled and finally, for each non-project PAs identified by HCS, 12 households (180/15) were sampled.

Households within each identified PAs were selected for inclusion in the sample frame through a process of consultation with PA committees applying the same criteria used for HCS participation, which requires households to be seed insecure for circumstances beyond their control and not related to being mediocre agricultural producers. Although this leads to potential self-section bias in the sample, it was essentially the only means by which sample frames at the peasant association level could be drawn up and the hope was to use econometric techniques to control for these sources of bias in any data analysis.

To practically select sample households, three lists of households randomly selected were provided by HCS:

List 1: HCS project participant households, *(from here on HCS- group)*
List 2: Non-project participant households in project PAs (*non-HCS1*)
List 3: Non-project participant households in non-project PAs *(non-HCS2)*

The method used to select households from each of the lists for sampling was to pick every tenth household until a total of 24 (12 for list 2 and 3) households had been selected. In cases where the end of the list was reached before attaining the required number, the selector had to start again with the top, leaving out those that had already been selected. When all required HHs for the sample had been selected, a list with their names and information on where they were located had to be made. As at times it may be difficult to locate households, we also compiled a group of alternate households. To get this, the procedure required to continue selecting every fifteenth household for 10 more households and make a list of these. These 10 would represent the alternate households in case it was not possible to survey some of the selected ones. The list of HCS-group would include 24 preferred households and 10 alternates. The list of non-HCS1 would include 12 preferred households and 10 alternates and so would the list of non-HCS2.

4.3.2 Community survey and community focus groups

The selection of communities for the community survey and focus groups was very straightforward. Since we wanted the community data to match the household and the community focus groups to provide insights that could be used for the data analysis, the sample for these included all the 30 selected PAs in which the household survey operated.
Leaders or PAs representatives from each PAs would be interviewed for the community survey.

A total of five focus group exercises were carried out. One focus group was held for Dire Dawa woreda, which consists only of mid-lands areas. Two focus groups, one representing the midlands one representing the highlands were held for each of the two remaining woredas, Chiro and Meta, for a total of four more focus groups. Each focus group was composed of a total of 15 men and women who had to be representative of each selected PAs from agro-ecological, ethnical, cultural and income level

aspects. Farmers were also requested to bring along samples of panicles or grain from their fields or harvest. The selection of representative farmers for each PAs was done by HCS in collaboration with leaders and/or representatives from each PA.

The survey instrument for community questionnaire can be seen in appendix 6.

4.3.3 Market survey

Relevant market outlets[17] for sampled HHs were identified as those that lay within one-hour walking distance of survey households. Applying these criteria, a total of eight markets serving surveyed areas were identified within which we then needed to select sellers to survey on a weekly basis. The sample selection required a representation of 100% of the sellers in case of small market outlets (up to a maximum of 8 sellers) and a representation of minimum 20% of the sellers in case of big market outlets (more than 40 sellers). In the latter case, the selection procedure required picking every third seller for sampling.

4.3.4 Agro-morphological survey

The sample selection for validating names and traits of varieties within the study area had to be consistent with the distribution of varieties most frequently used by surveyed HHs. Consequently the sample selected was a sub-group of HHs surveyed and the survey focused on the set of varieties for which focus group results suggested that names were consistently used.

In collaboration with IPGRI, IBCR, AU and HCS a set of 30-40 qualitative and semi-quantitative traits was defined for recording diversity and validating sorghum and wheat-variety names. The procedure involved evaluation of each of the selected varieties in 4 plots. Each plot had to be operated by a different HH. Varieties that were selected for validation were those recorded in the first round of HH survey and already validated through the focus groups.

Farmers sampled for the HH survey and who reported growing the selected varieties were grouped based their variety selection and these groups were then randomly sampled until 4 households were identified for the agro-morphological survey.

Furthermore, the agro-morphological survey was intended to provide insights on the impact of HCS activities on infra-crop diversity. To this purpose, information had to be collected about the sorghum and wheat varieties distributed by HCS. There were four sorghum varieties and two wheat varieties for which HCS impacts had to be checked. The sample procedure used was to sample plots from 10 randomly selected HCS participants and 10 non-HCS participants from comparable PAs. From each plot a total of 30 plants per variety had to be collected.

4.4 Data collection

Table 1 provides a summary of the different data collection instruments conducted as part of the field work. In this section, the details of each of the instruments used to collect the data are described.

[17] By "relevant" we meant a market place where the sampled HH would most commonly go.

Table 1. Description of surveys used for seed system analysis in Hararghe, Ethiopia

Instrument	Focus	Unit of Analysis	Number	Timing	Period of Focus
HOUSEHOLD SURVEY	Farmers seed use and selection behavior and household welfare/benefits	Households divided in 3 groups: HCS-group non-HCS1 non-HCS2	720	1st round: August 2002 2nd round January 2003	Cropping year 2002
COMMUNITY SURVEY	Community conditions which impact farmers' seed selection behavior	Peasant Association	30	1 time only	Cropping year 2002
COMMUNITY FOCUS GROUPS	Validation of variety names in sorghum and wheat; gender perspectives	Community focus groups	5	1 time only	October 2002
MARKET SURVEY	Market conditions for seeds and grains: prices and quantities by variety	Market outlets	7 markets	18 weeks	2002 harvest season: July 2002-January 2003
AGRO-MORPHOLOGICAL SURVEY	Agro-morphological characteristics of varieties farmers select	Farmer plots	30 varieties of sorghum 12 varieties of wheat 4 plots per variety	1 time only	October 2002 (meher season)

4.4.1 Household survey

The household survey focused on the cropping season of 2002. It was designed in such a way to collect direct information from farmers necessary to create key variables described in section 4.1.1 as well as other variables and instruments to be used as explanatory in the analysis to be conducted afterwards.

The first and second round of the HH survey questionnaires are respectively in Appendices 1 and 2.

A first draft of the questionnaire- prepared using information gathered through literature review, documents and details obtained throughout visits to Ethiopia - was tested during the first visit to Hararghe, necessary changes were made afterwards. Further amendments were made, incorporating comments from experts from HCS and Alemaya University. Particular attention was paid to the critical drought situation Ethiopia was facing in the moment the study was conducted as well as to the best way of gathering information regarding HIV/AIDS affecting the sampled area. We considered HIV/AIDS as an important topic within our survey because of the loss of labour and of agricultural biodiversity it can cause. We tried to combine input and suggestions received from the SDWP service of FAO on this issue, with input from CIAT and AGPP on seed emergency interventions and with comments and suggestions from the local teams based on the knowledge of the area. Moreover, comments from FAO experts on gender issues, seed breeding and commercialization as well as from experts from IPGRI and IFPRI on CGRs, and from University of Wageningen on the Hararghe Region of Ethiopia, were obtained and incorporated in order to address appropriately each of the issue relevant to the topic under study.
Before starting the actual data collection, the questionnaire was pre-tested throughout the enumerator training which combined classroom sessions with field exercises. Enumerators and team leaders' feedback as well as farmers' reaction to questions enabled the modification of some of the questions which were either irrelevant to the current situation or out of context

or difficult to understand. Furthermore, this process of continually revising and improving the survey instrument with the participation of the survey team was very positive in that, besides obviously enriching the questionnaire, it made the survey very much a team effort, thus increasing the enumerators' "ownership" and feeling of responsibility for the survey. The team leaders and enumerators proved to be competent and interested in the study and gave good input and suggestions. From the first to the last survey test, the time required was substantially reduced (from three hours to two hours). Major changes in survey design and implementation were required because of delayed planting caused by rainfall shortage.

In order to assure good quality data, a checklist to be completed and signed by each enumerator and his team leader on the last page of the survey was prepared. This had to be done right after the interview was completed and possibly before leaving the farmers so that eventual missing information could still be gathered and mistakes corrected.

As already pointed out the length of the household survey caused its split into two rounds for data collection. The first round was held towards the end of the Meher (main crop) planting season in August 2003. The second round was done after the harvest of the Meher crop in February 2003, delayed because of rain-delay. Consequently, during the second round of enumerator training, particular emphasis was given to the linkages between the first and the second round of the survey, on how to deal with incorrect pre-printed data and with modified situations in terms of family composition or operated land, on how to treat unavailability of HH members to be interviewed and impossibility to interview the same person within the household etc. Indeed, dividing the survey into two visits usually creates a number of problems mainly due to:

-sample reduction because of HH migration or unavailability at the time of the second round;
- different conditions in terms of operated land or family composition;
- mistaken or missing data gathered during the first round; etc.

During the last day of the training the FAO technical team gave instructions, recommendations and a list of common mistakes to avoid of daily survey management to enumerators and team leaders, distributed a copy of the code sheet and of enumerator guide in which detailed explanations on how to compile the questionnaire and purpose of each question, were distributed to each team member[18].

The AU team leaders were responsible for the quality of the surveys from each member of their teams. The team leaders had to sign off on the survey verifying the accuracy of the survey once it had been signed and submitted by their enumerators immediately after each survey had been completed. They also had to submit a list of completed surveys and submit the list to HCS. The surveys were then submitted to the HCS survey co-ordinator that could keep track of surveys done and of their level of accuracy verifying receipt of the surveys through his signature. By having everyone sign off on the completion of their tasks, the hope was that a high level of quality would be achieved and each of the respective partners would take responsibility for maintaining it.

Furthermore, to avoid conflicts between AU teams and HCS co-ordinators and in order to get good quality data, the FAO team agreed with the co-ordinator of the FAO-Norway funded

[18] Survey management, recommendations and enumerator guide are available in appendices 3, 4 and 5.

project based at Alemaya University to create and take the lead of a quality control board. The the quality control board was chaired by the FAO field representative, Dr. Osman, and included one representative from HCS and one from Alemaya University. The requested representatives could not be directly involved with the data collection efforts. The purpose of this board was to identify and resolve any potential problems arising during the data collection period that could affect the quantity and quality of data collection efforts, such as adverse weather conditions, lack of vehicle availability or breakdown, inability to locate sample farmers in a timely fashion, illness of enumerators or survey team leaders, poor performance of survey enumerators or team leaders or co-ordinators and so on. When any member of the data quality control board identified a problem, the board was requested to agree on discussing the problem and recommend measures to resolve it. The FAO field representative, in consultation with the FAO technical team in Rome would take the final decision on how to respond to the issues being raised.

Daily planning of the HH and community surveys was done jointly by HCS staffers and the AU team leaders. Thus HCS staff accompanied the surveyors at all times. A second HCS staffer was in charge of arranging for the vehicles and contracting the informants. Each survey team had to complete 12 household surveys per day for a total of 180 per team during the survey period (15 days) for both round 1 and round 2. Therefore, each enumerator should have completed about 45 surveys or 3 per day during the time of the survey and same procedure for the second round. Therefore a total of 720 households had to be surveyed twice. Since on the second round of the HH survey, one team was one surveyor short, because the missing enumerator had never joined the group as promised, 3 surveys a day had to be divided among the remaining enumerators of the same team.

Each day, one survey team had to go to the non-project PA and the other three teams to the project PA. Of the three teams that went to the project PA, one team would interview non-project participants and two teams would interview project participants. The teams were requested to rotate – that is, the same team should not go to the non-project PA everyday nor interview the participants. Team leaders were left to decide on the rotation keeping the survey coordinator informed of activities. The rotation for the second round worked on the basis of first round, which means each team interviewed the same groups of household interviewed last time. To do this, each team had a number assigned corresponding to the 4^{th} digit of each survey number. Team 1 households will all start with 1, team 2 will all start with 2, team 3 with 3 and team 4 with 4. The team leaders filled in the daily log writing the household survey number next to the name in the log. This system allowed us to checking for systematic mistakes once the data were inserted in the data base.

The day before the visit to the field, lists of 24 HCS-group and alternates, 12 non-HCS1 and alternates in HCS PAs, and 12 non_HCS2 in a non-project PAs and alternates were given to team leaders. The team leader assigned a questionnaire, and therefore a specified household, to each enumerator. The survey coordinator had to ensure that the households were available and willing to be surveyed. If that was not possible, the team leader provided the enumerator with an alternative household to survey from the alternate list for round one. As far as possible, the same person interviewed during the first visit had to be interviewed for the second round. However, if that person was not available, a direct substitute from the same household, such as his wife (or her husband) or one of the children, that can give answers instead of the head of HH was considered acceptable. In this case, the enumerator had to write a note in the cover page explaining why the same person as last time could not be interviewed. Similarly, if the enumerator could not find the person to interview nor another member of the same household a note explaining why the questionnaire was not completed had to be inserted on the cover page. The team leaders were responsible for making sure due notes were reported on the cover page in case a questionnaire was returned empty or questions were asked to a different member of the household.

The survey coordinator was in charge of managing all the forms and of providing them to team leaders for completion and for archiving and storing the forms once completed and signed. Survey coordinator was also responsible for checking the code numbers and for inserting the code on each page of the survey for the first round, while on the second round codes would be pre-printed. He would also keep track of all completed surveys and forms during the course of the survey.

At the same time of the first round, two community surveys were done each day: one community survey in the project PA and one in the non-project PA. The coordinator had to arrange for leaders or representatives from the communities to be interviewed. The community surveys were done by two of the team leaders each day so each will have done 7 or 8 by the end of the survey. The team leader going to the non-project PA with his team was asked to survey that PA. The team leader going to the project PA with his team and surveys the non-project households would do the community survey for that PA. Thirty PAs have been visited in 15 days – 15 project PAs and 15 non-project PAs. Therefore, each day one project PA and one non-project PA have been visited.

Despite the many problems encountered, survey management seemed to work quite well during the first week of the survey under the FAO technical team supervision. Most of the mistakes were negligible and easy to correct by checking the surveys on the day the data was collected. However, to avoid mistakes and bad quality data, after the first day monitoring, a list of recommendations was compiled and was distributed to team leaders with explanations on things to pay particular attention to. The biggest problem encountered, for the first round of the survey, was the strategic behavior of farmers who were hoping to gain some rewards based on the way they responded to the questions (e.g. get seed aid if they reported no seed in stock etc.). To this purpose recommendations were given to double check questions with farmers in order to elicit the real answer. Conversely, the biggest problems encountered during the second visit were linked to incorrect data reported in the first visit. Careful attention was recommended to be paid to this problem, also because most of the mistakes reported from the first visit were clearly due to lack of attention from team-leaders and/or from enumerators.

4.4.2 Community survey

The survey of community characteristics in all of the PAs where sampled households reside was conducted at the same time as the first round of the household survey. The PA survey served to complement the HH survey with data on seed system and community characteristics from another scale. For each of the selected PA a survey of community characteristics regarding population, main religious and ethnic groups, infrastructure, marketing development, market facilities and sources of seed supply were collected. A key informant from the community, usually a government official, was the source of information for this survey which was completed in each PA by two of the team leaders on a rotational basis. During the enumerator training, we also gave a separate training to the team leaders for conducting the community survey which was much shorter and easier as compared to the HH survey. The last three days of field test for the 1st round of the HH survey, the team leaders tested the community survey as well as the daily management of survey activities under supervision of FAO team.

Selected communities were also geo-referenced in order to allow spatial analysis and therefore control for spatial distribution of the impacts the seed system intervention is likely to have.

4.4.3 Community focus groups

The focus groups were implemented after the first round of the household survey but before the agro-morphological characterization of sorghum and wheat varieties.

Once representatives of PAs to include in the five focus groups were selected, as described in section 4.3.2, meetings of half day were organised with the help of six facilitators, including both men and women, who were in charge of recording a common agreed set of mechanisms and procedures.

Prior to actual implementation of focus groups, a 2-day workshop was organised with the facilitators in order to train them on how to elicit and compile information on farmers' characterization of landraces as well as on how to enumerate and compile farmer's information in a simple spreadsheet that would allow the landrace name and its associated information to be correlated with the socio-economic data. Training included practical exercises with theory given beforehand. Participatory discussion was held on specific cases locally selected. To give supplementary insights a number of related articles and publications were provided. Ways of preparing social and resource maps of a village along with its significance in R&D programs was discussed.

The second section of the program consisted of the field exercise. Once the preparatory work for the field exercise was over, the team developed a checklist. The checklist was discussed with the team for clarification and also to re-orient the questions where necessary. To make all project staff and farmers familiar in terms of the social composition, infrastructure and resource distribution within the project areas, social maps involving key farmers was prepared. Overlaying altitude, latitude, longitude and temperature on these maps as a reference point through GPS reading provided a useful guide. The results of the focus group proved to be rather successful. Moreover, farmers demonstrated appreciation of the variety of information acquired through focus group. The crucial issue to the success of the study was to have been able to find the right individuals for understanding the institutional setting.

4.4.4 Market survey

The market survey involved the weekly collection of data over a two month period (at planting time) from local markets where seed and grain were sold. The markets were largely located within one hour distance from sampled Peasant Associations, or representative markets were selected where farmers would most commonly go. Data was collected by HCS designated staff on the prices and quantity of identified sorghum and wheat varieties on a weekly basis from sellers selected as described in section 4.3.3. For each variety, prices and quantities for seed and grain sold would have been recorded and, within seed and grain a further distinction by quality level had to be recorded whenever possible. Varieties were identified on the basis of pre-survey visits, to selected markets as well as from first results from HH survey.

Implementation of market survey was rather smooth, although poor weather conditions prevented data collection over some of the assigned collection periods. Data were inserted in the spreadsheet in such a way to require total re-inputting or check for each of the data and market inserted. One major problem in the survey implementation was the difficulty in getting information on prices that distinguished seed from grain quality and the lack of information on variety specific seed. In most cases seed sold in local markets was bulked into a generic variety class such as "red wheat" which consisted of a blend of several varieties. According to local experts, small traders (often local women selling on the margins of the marketplace) were more likely to have seed distinguished by variety, but these were not covered by the survey.

4.4.5 Agro-morphological survey

The main objective of the survey was to analyze the content of variety names, the degree to which they are consistent across and within communities as well as the extent and distribution of diversity of sorghum and wheat within the study area. The range of plant types, maturity periods, disease and pest resistance characteristics, grain properties would have been elicited via the survey indicating how much diversity farmers' use within a village, PA or Woreda. The survey required characterization of 30 plants per plot in farmers fields, listing all traits of a given variety.

The study thus involved the collection of data on the agromorphological characteristics of the varieties supplied by HCS and comparing these with the same varieties from other sources. Analysis of numbers and distributions of local varieties in HCS and non-HCS areas was also undertaken. Data collection had to be limited to traits of high heritability not subject to significant GxE interactions. Training for the field enumerators was provided by a designated staff from IBCR over a 3 day period in November 2003.

Prior to the implementation of the agro-morphological survey, a training workshop lead by IBCR for data collection and reporting was held at HCS over a 3 day period in November. Expectations were that the agro-morphological data collection would take about 2 h. per plot and therefore require 32 person days for sorghum and 10 person days for wheat (approx 10 enumerators for 5 days). Unfortunately, the late rain combined with prior drought required a postponement of the data collection and complicated the selection of HH as for most of them no samples of grain were still available in the fields.
However, fourteen farmer varieties of sorghum and twelve of wheat from the three study

woredas were selected. For each farmer variety of the two crops 30 individual plants were used to record 27 traits for sorghum and 20 for wheat. Characters were grouped into quantitative and qualitative as described in section 4.3.4. The former group was analyzed using statistical approach with the support of SPSS software. The qualitative characteristics were used to create a Simpson diversity index which measures the *evenness* of the distribution of traits across woredas and altitude ranges. Mr. Tanto was in charge of checking for data entry and data quality with general supervision provided by AU and IPGRI in order to allow for a structured analysis of variance and for multivariate analysis techniques to be linked to HH survey data.

4.5 Data entry

The program used to enter HH and community surveys data is the Integrated Microcomputer Programming System that has been developed by the Central Bureau of Statistics for data entry and consistency checks. It is a question-oriented program, which means it has the same format of the questionnaire used to collect data. After data were entered, a consistency check program called Concur to double check for mistakes and for consistency was used. Once the second program was executed, each table was saved in a different file into SPSS format, which was easy to import in Stata, the program we used for econometric and statistical analysis.

The person in charge of programming the software is an expert from the central bureau of statistics -Mr. Tabit Amhed- to whom final version of survey instruments (HHs and PAs), code sheets and enumerator guides were provided. Moreover we provided also a list with most common mistakes and, recommendations to enumerators and team leaders, which, together with the checklist at the end of the HH survey, should serve as a good quality control system to start up with designing the control program.

The entry program was designed in such a way to take care of:

- Avoiding duplicate entry of surveys
- Linking data from first visit to second visit
- Setting appropriate range limits for variables (e.g. refusing responses higher than allowed on survey)
- Setting missing values and "don't know" answers as 99, 999, 9999 etc depending on the number of digits for the question into consideration
- Labelling each code in order to link numbers to names.

The manual imputation of data was the responsibility of HCS who designated a couple of staff members to doing it while the person who had co-ordinated the survey process for both first and second rounds, was requested to supervise the process in order to have as a supervisor someone who knew the survey instruments as well as the way data were collected.

With regard to the market survey, the data were collected in forms set up as the spreadsheets in which data were thereafter entered. This process was also responsibility of HCS.

Finally IBCR and AU were in charge of data checking and results-review of agro-morphological and focus groups with support from IPGRI. AU and IBCR were responsible for overseeing analysis of data to provide means and variances for the varieties and over the

different levels of study (PA, Woreda, study area). Consistency estimates for variety names and key traits had to be calculated, using established procedures, by HCS designated staff with support from IBCR, AU, IPGRI and the focus group trainer. The data analysis from these components of the study would then be combined with the data analysis of variety number and distribution from the overall survey once this had been corrected for non-consistent variety names. Excel spread sheets were used for data entry with SPSS analysis packages. IBCR and AU took the lead in this work with inputs from IPGRI as agreed.

5. LESSONS LEARNED

5.1 Key implementation issues

A serious drought was experienced in Hararghe in the 2002 cropping year, resulting in high rates of crop failure and increased rates of food insecurity and hunger in the areas where the study was conducted. This situation presented several difficulties in the implementation of the work, but also created an important opportunity to gain insight into coping mechanisms (or lack thereof) in the seed system. Moral issues arise in using resources to study people's suffering – rather than to alleviate it. Obviously it is important to understand what happens in the event of a crisis such as a drought in order to design effective strategies to alleviate human suffering, but it is hard to be in the position of implementing a study under these conditions. It is extremely difficult for a well-fed and relatively wealthy person to ask a poor farmer how many days they have had to go without food. The survey enumerators in the second round clearly felt this difficulty. They reported that the farmers thought they were there to provide relief food or seed and in some cases became angry and aggressive with the survey teams when they realized it was not a relief operation. The FAO technical team consulted with the Alemaya researchers and HCS staff about how to handle this problem and we decided that farmers would be paid for participating in the second round of the survey, despite the fact that this is considered to be poor survey procedure, resulting in biased data. In this case we were surveying farmers which had already participated in the first round of the survey for free and paying them the equivalent of what they could have earned (mostly from collecting and selling fuelwood) during the time it took to respond to the second round of the survey, and this seemed imminently reasonable under the circumstances.

Another impact of the drought and crop failure was problems in collecting data for the agro-morphological study; the crops were not in the field to be evaluated at the time of the survey. For this reason, the objectives of the agro-morphological study were not fully realized and we could not obtain the sample of varieties that were desired. In some PAs, none of the surveyed farmers had crops in the field to evaluate. In some cases the agro-morphological survey team evaluated varieties from other farmers in the community who were reported to be growing the same varieties as the surveyed farmers. The sample obtained in the agro-morphological survey was therefore not sufficient to allow us to evaluate the impacts of the HCS program on infra-specific diversity, however the data was analyzed and did provide insights on the distribution of various measures of crop diversity within and among communities in the study. The data could have been even more useful if the original plan for the sample to be linked to the household survey data could have been achieved.

5.2 Lessons learned and recommendations

Some important lessons learned have emerged from this experience on how to approach the analysis of seed supply on farm levels of diversity and welfare.

One key question is the balance and timing of quantitative vs. qualitative data. In this effort, the community focus group surveys proved to be very rich sources of information that would have been useful in designing the household survey and sample. Since they were conducted after the first round of household surveys however, we were not able use this information as well as would have been the case, had it been collected prior to the first round of the household survey.

Recommendation: Implementing qualitative studies, such as community focus group studies prior to any formal quantitative data collection is highly recommended in order to ensure the integrity of survey and sample designs.

Another issue which arose in this study is the validity of the household survey data on seed supplies and sources, due to the presence of several NGO and government programs to distribute seed in the area, which could have lead farmers to report more problems with seed sourcing than is actually the case. Problems with moral hazard arise in these situations which are very difficult to deal with.

Recommendation: In situations where moral hazards exist in terms of farmer response to questions (e.g. on seed supply quantity and sources) care needs to be taken in how to elicit reliable data. One way is to structure questions to ask about overall seed quantities planted, then the amounts obtained from various off-farm sources, and from this estimate saved seed quantity. Another method is to build in checks in the survey; asking the same question in more than one way, to allow for some verification. Alternatively, acquiring data from neighbors and/or key informants on seed supply sources and quantities used by farmers in the area could be used as a means of verification. An even more effective means could be the monitoring of farm HHs for one full year by a village resident who will be given a reasonable incentive and supervised on a weekly basis by a supervisor.

An important problem facing studies of seed systems and genetic diversity is developing a means of comparing the genetic and phenotypic diversity present at varying scales in the system, given a lack of comparability in variety names among locations and actors in the seed system, as well as different practices of aggregating and mixing seed. In this study, the market data have not proven very useful in analyzing the household varietal choice, because different definitions, naming and aggregations of varieties at the market level make it difficult to relate to household varieties.

Recommendation: This is an issue that focus groups and qualitative interviews with various actors in the seed system may help to resolve. At a minimum, the collection of qualitative information about the relationship between variety names and attributes and how they vary among households, communities and within components of the seed system will help identify the degree to which variety names can be used as a comparable unit of analysis. Other means may be the use of agro-morphological analysis and variety grow outs from material obtained at various points in the seed system, as well as molecular analysis. Many landraces are grown as a mixture of varieties and characterizing these as one variety is misleading. A grow out test may allow for the quantification of the proportion of such varieties grown in mixtures but obscured by the 'major' variety. Farmers should be the ones who identify such varieties in a grow out test, in collaboration with breeders who make use of the agro-morphological information documented for the landrace collections at IBCR. Until such empirical relationship are established between information obtained from farmers and from grow out and molecular marker studies, one may not realistically capture the degree of diversity.

Splitting the survey into 2 rounds is a very delicate and tricky exercise which can easily lead to some problems and mistakes due to sample reduction, different demographic and land owning conditions, mistakes or missing data gathered during the first round.

Recommendation: To the extent possible, multiple rounds of data collection in a household sample survey should refer to unrelated and different sections of the household (i.e. socio economic for one round and ag. production for the other). Otherwise, very careful attention

should be paid to gather precise information asking questions in the easiest way possible avoiding any kind of ambiguity: for example allowing the enumerators to gather data using their own unit of measurement, code answers as much as possible etc. Also, it would be recommendable to have the data set from first round up and running at the time of the second survey, which would require receiving data as early as possible.

Problems with data entry errors greatly slowed the progress of the data analysis and also lead to lower data quality.

Recommendation: More attention should have been paid at data entry level with close supervision provided by a senior member of the survey team who understood the issues involved in each question and its responses.

While most enumerators proved to be quite responsible and motivated, closer supervision was needed in order to ensure good data quality. Supervision was also needed because of the difficulties created by having different counterparts and with the consequent complexity of having clear responsibilities.

Recommendation: If possible one single counterpart should be contracted and considered in charge of implementing the project. Provided this, careful attention should, therefore, be paid in selecting team leaders that are responsible and not too senior to accepting a role which involves also going out to the field and adapt to, sometimes, harsh conditions and to be still enough motivated in doing this kind of exercise. In addition, it is important to have a person from the technical team that designed the survey (e.g in this case FAO) follow the entire process of data collection in the field in order to ensure data quality.

A tedious and vast amount of work has been required for cleaning data of the HHs survey. Many of the problems were related to linking first with second round of data collection, which required big efforts and precision to match 1st round with second round and obviously involved the dropping out of many surveys. Indeed, most of the mistakes were not systematic and therefore required an evaluation and careful checking of mismatching on a survey by survey basis. Some mistakes were due to the conversion of unit of measurement (from Ethiopian measures to western ones and vice versa) which could have been avoided if we had requested the enumerators to keep their units so that we would convert them after receiving data. Some other mistakes were made at data input level and could have been easily avoided with a little more supervision. However, despite the work required for data cleaning which prolonged unnecessarily the time needed for setting the data up and running, both HH and community surveys formed a good data set to be used for analysis in conjunction with focus group that also provided very good insights and secondary information.

Another problem was due to the insufficient variation in socio-economic conditions. Most of the farmers interviewed had a very similar level of poverty which created some difficulties in capturing differences possibly attributable to a different level of wellbeing. Last but not least, it was difficult to capture any differentiation between seed and grain in the markets and also in the quality levels impacting the price. Farm level distinctions between varieties and seed quality were generally not reflected at the market level and a key problem with the market data is an inability to relate it back to what farmers reported planting. This is an important issue that must be better addressed in future studies that incorporate data collection on seeds and varieties at more than one scale of analysis.

Recommendation: in future studies, more attention should be paid to these differences in socio-economic conditions as well as in variety names and aggregation is needed in order to trace the impacts of seed supply on farm crop biodiversity.

References

Abbott, J. A., (2005), Counting Beans: Agrobiodiversity, Indigeneity,and Agrarian Reform, *The Professional Geographer,* 57(2) :198–212

Adenew, B., H., Storck and M., Mulat, (1991), *The analysis of Land Size Variation and its effects: The case of Smallholder farmers in the Hararghe Highlands*, Studies on the Smallholder Agriculture in the Hararghe Highlands, Eastern Ethiopia, Working Paper No. 66, Institute of Horticulture Economics, Hannover University, Germany.

Aguirre-Gomez, J. A., M. R. Bellon, and M. Smale, (2000), A regional analysis of maize biological diversity in southeastern Guanajuato, Mexico, *Economic Botany,* 54(1): 60-72.

Alemayehu, F., (1995), *Genetic variation between and within Ethiopian balrey landraces with emphasis on durable resistance*, PhD dissertation, University Wageningen, Landbouw

Almekinders,C. and D. Louette, (2000), Examples of Innovations in local seed system in Mesoamerica, in C. Almekinders and W. de Boef (eds.) *Encouraging Diversity*, Intermediate Technology Publications London

Arnason, J.T., J.D.H. Lambert, J. Gale, J. Mihan, M. Bjarnason, D. Jewell, J. A. Serratos, J. Fregeau-Reid, and L Pietrzak, (1993), Is "quality protein" maize more susceptible than normal cultivars to attack by the maize weevil Sitophilus zeamais?, *Post harvest Biology and Technology,* 2: 349-358.

Ayana, A., and Bekele, E., (1998), Geographical patterns of morphological variation of sorghum (sorghum bicolour) (L Moech) germplasm from Ethiopia and Eritrea: qualitative characters, *Hereditas*, 129(3): 195- 205

Badstue, L. B., (2004), Identifying the Factors that Influence Small-scale Farmers' Transaction Costs in Relation to Seed Acquisition An ethnographic case study of maize growing smallholders in the Central Valleys of Oaxaca, Mexico, FAO-ESA working paper, 04-16.

Baniya B.K., Subedi A., Rana R.B., Paudel C.L. Khatiwada S.P., Rijal D.K., Sthapit B.R. (2000) Nepal. Seed supply systems: data collection and analysis. In: Jarvis D., Sthapit B., Sears L. *Conserving agricultural biodiversity in situ: A scientific basis for sustainable agriculture. International Plant Genetic resources Institute*, Rome Italy. pp. 159-164

Baumgärtner, S. (2004). Measuring the diversity of what? And for what purpose? A conceptual comparison of ecological and economic measures of biodiversity, *Working Paper*, Department of Economics, University of Heidelberg, Germany

Bechere, E., Kebede, H., and Belay, G., (2000), *Durum Wheat in Ethiopia, An Old Crop in an Ancient Land,* Institute of Biodiversity Conservation and Research, Addis Ababa: Ethiopia.

Bellon, M. R. (1996a), Landholding fragmentation: Are folk soil taxonomy and equity important? A case study from Mexico, *Human Ecology,* 24(3): 373-393.

Bellon, M. R., (1996b), The Dynamics of Crop Intraspecific Diversity: a conceptual Framework at the Farmer Level, *Economic Botany,* 50(1), 26-39.

Bellon, M.R. (2004), Conceptualizing Interventions to Support On-Farm Genetic Resource Conservation, *World Development:* 32(1): 159-172.

Bellon, M.R. and Risopoulos, J. (2001) Small-scale farmers expand the benefits of maize germplasm: A case study from Chiapas, Mexico. *World Development* 29 (5), 799-812.

Bellon, M.R., (1997), On-Farm Conservation as a Process: An Analysis of Its Components, In:, L. Sperling and M. Loevinsohn(eds.), *Using Diversity, Enhancing and Maintaining Genetic Resources On-farm,* International Development Research, Canada.

Benin, S., B. Gebremedhin, M. Smale, J. Pender and S. Ehui (2003) Determinants of Cereal Diversity in Communities and on Household Farms of the Northern Ethiopian Highlands, *ESA Working paper,* No 03-14, Agricultural and Development Economics Division, FAO, Rome: http://www.fao.org/es/ESA/pdf/wp/ESAWP03_14.pdf

Beyene, H., Verkuijl, H. and Mwangi, W. (1998) *Farmers' Seed Sources and Management of Bread Wheat in Wolmera Woreda*, Ethiopia. CIMMYT and IAR, Mexico D.F.

Bramel, P., Nagado, J.S., Haugegen, M. J., Adugna, D., Dejene, T., Bekele, T., Traedal, T. D., (2004), Relief Seed Assistance in Ethiopia, in: L. Sperling, T. Remington, J. Haugen and S. Nagoda (eds.), *Addressing Seed Security in Disaster Response: Linking Relief with Development*, Cali, Colombia International Center for Tropical Agriculture.

Brush, S. B., (1995), In Situ Conservation of Landraces in Centers of Crop Diversity, *Crop Science,* 35(2): 346-354.

Brush, S.B., 1995, In situ conservation of Landraces in Centers of Crop Diversity, *Crop Science*, 35(2): 346-354

Brush, S.B., J.E.Taylor and M.R. Bellon, (1992), Biological diversity and technology adoption in Andean potato agriculture. *Journal of Development Economics,* 39(2):365-387.

Byerlee, D. and P. W. Heisey, (1996), Past and Potential Impacts of maize research in sub-Saharan Africa: a critical assessment, *Food Policy***,** 21(3): 255-277.

Byerlee, D., and P.W. Heisey, (1997), Evolution of the African Maize Economy, In D. Byerlee and C.K. Eicher (eds.), *Africa's Emerging Maize Revolution.* Boulder, Colorado: Lynne Rienner Publishers.

Central Statistics Authority (CSA). 1999. Agricultural sample survey (1998/99): Report on area and production for major crops for private peasant holding, main season. *Statistical Bulletin 200.* Addis Ababa, Ethiopia: CSA.

Convention on Biological Diversity: http://www.biodiv.org/convention/articles.asp

Cromwell, E., Friis-Hansen, E. and Turner, M., (1992), The seed sector in developing countries: A framework for performance analysis, *ODI Working Paper* 65

de Janvry, A., Fafchamps, M., and Sadoulet, E. (1991) Peasant household behavior with missing markets: Some paradoxes explained. *The Economic Journal* 101:1400-1417.

Dercon, S. (2001), Economic Reform, Growth and the Poor: Evidence from Rural Ethiopia, *Working paper n. 8, Centre for the Study of African Economies*, Dept of Economics and Jesus College, Oxford University

Dercon, Stephan, 1993. The consequences of liberalisation and peace for food markets in Ethiopia, unpublished.

Douglas, J.E. 1980. *Successful Seed Programs: A Planning and Management Guide.* Boulder, Colorado: Westview Press.

FAO: Food and Agriculture Organization of the United Nations (2000). The 2000/2001 Joint FAO/WFP crop and food supply assessment mission to Ethiopia, FAO mimeo.

FAO-GIEWS, FAO- Global Information Early Warning System, (2005), *Special Report Crop and Food supply Assessment Ethiopia*, ESC/Reports, FAO- Rome, Italy, http://www.fao.org/docrep/007/j3958e/j3958e00.htm

FAOSTAT, (2004), http://faostat.fao.org/faostat/default.jsp?version=int&hasbulk=1

Gabre-Madhin, Eleni Z. 2001. *Market Institutions, Transaction Costs, and Social Capital in the Ethiopian Grain Market.* IFPRI Research Report 124, International Food Policy Research Institute, Washington, D.C.

Geleta, M., Z. Asfaw, E. Bekele and A. Teshome, (2002), Edible oil crops and their integration with the major cereals in North Shewa and South Welo, Central Highlands of Ethiopia: an ethnobotanical perspective, *Hereditas,* 137:1

Gemeda, A., Aboma, G., Verkuijl, H., and Mwangi, W., (2001), *Farmers' Maize Seed Systems in Western Oromia, Ethiopia.* Mexico, D.F.: International Maize and Wheat Improvement Center (CIMMYT) and Ethiopian Agricultural Research Organization (EARO).

Godoy, R., V. Flores, D. Bravo, P. Kostishack, A. Cubas, K. O'Neill, K. McSweeney, and D. Wilke, (2000), Human capital, wealth, property rights, and the adoption of new farm technologies: The TawahkaIndians of Honduras, *Human Organization,* 59(2): 222–33.

Hailu Gebremariam. 1992. *Availability and Use of Seed in Ethiopia.* Addis Ababa, Ethiopia: Program support unit, Canadian International Development Agency.

Hailu, G.M., (1991), Wheat production and research in Ethiopia. In: Hailu Gebre-Mariam, D.G. Tanner, and M. Hulluka (eds.). *Wheat Research in Ethiopia: A Historical Perspective*, IAR/CIMMYT: Addis Ababa.

Harlan, J.R., (1969), Ethiopia: a center of diversity, *Economic Botany,* 23: 309-314.

Hassan, R.M., M. Mekuria, and W. Mwangi, (2001), *Maize Breeding Research in Eastern and Southern Africa: Current Status and Impacts of Past Investments Made by the Public and Private Sectors 1966-97*, Mexico, D.F.: CIMMYT.

Heisey, P. W., M. Smale, et al. (1997). "Wheat rusts and the cost of genetic diversity in the Punjab of Pakistan." American Journal of Agricultural Economics **79**(3): 726-737.

Heywood, V. 1999. Trends in agricultural biodiversity. p. 2–14. In: J. Janick (ed.), *Perspectives on new crops and new uses*. ASHS Press, Alexandria, VA.

http://www.cia.gov/cia/publications/factbook/docs/notesanddefs.html#2119
http://www.ers.usda.gov/publications/GFA16/GFA16fm.pdf.

ICCPPGR/FAO (1997), *FAO: Plant Genetic Resources*, FAO-Rome, Italy, http://web.icppgr.fao.org/

Institute of Agricultural Research (IAR). 1996. *Progress report for the period April 1995 to March 1996.* Addis Ababa, Ethiopia: Holetta/IAR.

IPGRI, 1997: Jarvis, D. I. and Hodgkin, T. (eds), Strengthening the scientific basis of is-situ conservation of agricultural biodiversity on-farm. Options for data collecting and analysis, IPGRI, Rome.

Jarvis, D. L., L. Myer, H. Klemick, L. Guarino, M. Smale, A.H.D. Brown, M. Sadiki, B. Sthapit and T. Hodgkin (2000). A Training Guide for In Situ Conservation, IPGRI. Rome,

Kebebew, F., Y. Tsehaye and T., McNeilly, (2001), Morphological and farmers cognitive diversity of barley (hordeum vulgare [L. Poaceae]) at Bale and north Shewa of Ethiopia, *Genetic Rouserces and Crop Evolution,* 48: 467-481

Key, N., Sadoulet, E., and A. deJanvry, (2000), Transaction Costs and Agricultural Household Supply Response, *American Journal of Agricultural Economics,* 82(2): 245-259.

King, A., (2000), A brief review of participatory approaches to use and conservation of agrobiodiversity. In: E Friis-Hansen and B Sthapit (eds.) *Participatory approaches to the conservation and use of plant genetic resources*. IPGRI, pp. 27-4

Lakew B., Semeane Y., Alemayehu F., Gebre H., Grando S., van Leur J.A.G., and Ceccarelli S., (1997), Exploiting the diversity of barley landraces in Ethiopia, *Genetic Resources and Crop Evolution*, 44(2): 109-116

Lirenso, Alemayehu, 1987. Grain marketing and pricing in Ethiopia: a study of the impact of grain quota and fixed grain prices on grain producers, *Research Report* No. 28, Institute of Development Research, Addis Ababa University

Longley, C.A., (1997) *Effects of War and Displacement on Local Seed Systems in Sierra Leone.* In L. Sperling (Ed.)- AgREN Network Paper No. 75. July 1997.

Louette, D. and M. Smale (2000), Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico, *Euphytica,* 113(1): 25-41.

Louette, D., Charrier, A., and Berthaud, J., (1997), In situ conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community, *Economic Botany*, 51(1): 20-38.

Louwaars, N., (1994), *Seed supply systems in the tropics: international course on seed production and seed technology*, Wageningen, The Netherlands: International Agriculture Centre.

MacDonald, K. (1998), Rationality, representation, and the risk mediating characteristics of a Karakoram mountain farming system, *Human Ecology* 26(2): 287–321.

Magurran, A. (1988) Ecological diversity and its measurement. Princeton University Press, Princeton, New Jersey, USA.

McGuire, S. (1999), Farmers' management of sorghum diversity in eastern Ethiopia, In C. Almekinders and W. de Boef (Eds.) *Encouraging Diversity: the conservation and development of plant genetic resources*, London: Intermediate Technology Publications.

McGuire, S. (2005), *Getting Genes: Rethinking seed system analysis and reform for sorghum in Ethiopia*, Wageningen University

MEDAC, Ministry of Economic Development and Cooperation, (1999), *Survey of the Ethiopian Economy: A Review of Post -Reform Developments*, Addis Ababa

Meng, E. C.-H. (1997), Land allocation decisions and In situ conservation of crop genetic resources : the case of wheat landraces in Turkey**:** 208 leaves. Ph.D. dissertation. Davis, California, USA: University of California

Morris, M.L. and Heisey, P.W. (1998) Achieving Desirable Levels of Crop Diversity in Farmers' Fields: Factors Affecting the Production and Use of Commercial Seed. In: Smale, M. (ed.) *Farmers, Gene Banks and Crop Breeding: Economic Analyses of Diversity in Wheat*, *Maize and Rice*. Klower Academic Press, Norwell, Massachusetts, USA.

Mulatu, E. (2000) Seed Systems and Small-Scale Farmers: A Case Study of Ethiopia. PhD Thesis. University of the Free State, South Africa.

Mulatu, E., and Kassa, H., (2001) Evolution of Smallholder Mixed Farming Systems in the Harar Highlands of Ethiopia: The Shift Towards Trees and Shrubs**,** *Journal of Sustainable Agriculture,* 18(4): 81-112

Mulatu, E., Osman, E. I., and Bekele, E., (2005), Policy Changes to Improve Vegetable Production and Seed Supply in Hararghe, Eastern Ethiopia, *Journal of Vegetable Science*, 11(2): 81-106.

Mwangi, W. and H. Verkuijl, (1998), *Natural Resource Policy Analysis*, - Teaching notes for the training workshop on Natural Resource Policy Analysis, CIMMYT, Addis Ababa, Ethiopia.

Negassa, M., (1985), Pattern of phenotypic diversity in an Ethiopian barley collection and the Arusi-Bale highlands as a centre of origin of barley, *Hereditas,* 102: 139-150

NSIA: National Seed Industry Agency (2003) Crop Variety Register, Issue n. 5, NSIA, Ethiopia.

Osborne, T., (2005), Imperfect competition in agricultural markets: evidence from Ethiopia, Journal of Development Economics, 76: 405–428

Rana, R., Shrestha, P., Rijal, D., Subedi, A., and Sthapit, B., (2000), Understanding farmers knowledge systems and decision-making: participatory techniques for rapid biodiversity assessment and intensive data plots in Nepal, In: E Friis-Hansen and B Sthapit (eds.) *Participatory approaches to the conservation and use of plant genetic resources*. IPGRI, 2000, pp. 117-126.

Rausser, G.C. and Small, A.A., (2000), Valuing Research Leads: Bioprospecting and the Conservation of Genetic Resources, *Journal of Political Economy,*.(1):173-206.

Refera, A. (1999), Tef: Post harvest Operations, In: *Compendium on Post-harvest Operations* (1999), Mejia, D. and B. Lewis (eds), FAO, Rome

Remington, T., J. Maroko, S. Walsh, P. Omanga and E. Charles (2002), Getting of the Seed and Tools treadmill with CRS seed vouchers and fairs. *Disasters* 6(4): 302-315.

Rice, E., Smale, M. and Blanco, J.-L. (1998), Farmers' Use of Improved Seed Selection Practices in Mexican Maize: Evidence and Issues from the Sierra de Santa Marta, *World Development* 26 (9), 1625-1640

Richards, P., (1986), Coping with hunger: hazard and experiment in an African rice-farming system, Allen and Unwin, London.

Seyfu, K. (1997). Tef, *Eragrostis tef* (Zucc.) Trotter. *Promoting the conservation and use of underutilized and neglected crops.12*. International Plant Genetics Resources Institute (IPGRI), Biodiversity Institute, Addis Ababa, Ethiopia. pp.50

Shiferaw, B. and Holden, S. T., (1997), A Farm Household Analysis of Resource Use and Conservation Decisions of Smallholders: An Application to Highland Farmers in Ethiopia, *Discussion Paper* #D-03/1997, Department of Economics and Social Sciences, Agricultural University of Norway (AUN), Ås.

Shiferaw, B., and Holden, S., (1999), Soil erosion and smallholders'conservation decision in the highlands of Ethiopia, *World Development*, 27:739-752.

Simpson, R. and D. Sedjo, (1998), The Value of Genetic Resources for use in Agricultural Improvement, in *Agricultural Values of Plant Genetic Resources*, R. Evenson, D. Gollin and V. Santaniello (eds.), Wallingford, CABI.

Simpson, R. D., R. A. Sedjo and J.W. Reid, (1996), Valuing Biodiversity For Use in Pharmaceutical Research, *Journal of Political Economy*,104(1): 163-185.

Smale, M. (1998). Farmers, gene banks and crop breeding : economic analyses of diversity in wheat, maize, and rice. Bostons, Kluwer Academic.

Smale, M. and International Maize and Wheat Improvement Center (1999). Farmer management of maize diversity in the central valleys of Oaxaca, Mexico : CIMMYT INIFAP 1998 baseline socioeconomic survey. Méxoco, D.F., México, Cimmyt.

Smale, M. and M.R. Bellon. (1999). A conceptual framework for valuing on-farm genetic resources. In D. Wood and J. Lenne (eds.), Agrobiodiversity: Characterization, Utilization, and Management. Wallingford, CAB International

Smale, M., Bellon, M. and Aguirre Gomez, J. (2001) Maize Diversity, Variety Attributes, and Farmers' Choices in Southeastern Guanajuato, Mexico. *Economic Development and Cultural Change* 50 (1): 201-225

Smale, M., R. P. Singh, et al. (1998). "Estimating the economic impact of breeding nonspecific resistance to leaf rust in modern bread wheats." Plant Disease **82**(9): 1055-1061.

Solow, A. and Polasky, S., (1994), Measuring biological diversity, *Environmental and Ecological Statistics* 1(2): 95-107.

Solow, A., Polasky, S., and Broadus, J., et al. (1993). On the measurement of biological diversity, *Journal of Environmental Economics and Management,* 24(1): 60-68

Sperling, L. and Cooper, D., (2003), *Understanding seed systems and strengthening seed security*. FAO, Rome: internal document.

Storck, H., B. Emana, B. Adenew, A. Borowiecki and S.W. Hawariat, (1991), *Farming systems and farm management practices of smallholders in the Hararghe Highlands – A baseline survey-,* Farming System and resource economics in the tropics, Vol. II, Wissenschaftsverlag vauk Kiel: Germany.

Tafesse, K., (2002), *Towards Seed Industry Development in Ethiopia,* FAO- Rome Italy, http://www.fao.org/ag/agp/agps/georgof/Georgo17.htm#Ethiopia

Takasaki, Y., B. Barham, and O. Coomes. (2000), Rapid rural appraisal in humid tropical forests: An asset possession-based approach and validation methods for wealth assessment among forest peasant households, World Development 28(11): 1961–77.

Tanto, T. and A. Demissie, (2000), *A comparative study of genetic diversity of four major crops managed of Ethiopia,* Institue of biodiversity Conservation and research, Addis Ababa.

Teshome, A. (2001) Spatio-Temporal Dynamics of Crop Genetic Diversity and Farmers' selections in-situ, Ethiopia. PhD Dissertation. University of Ottawa, Canada..

Teshome, A., J.K. Torrance, B. Baum, L. Fahrig, J.D.H. Lambert and J. T. Arnason, (1999), Traditional Farmers'Knowledge of Sorghum (Sorghum bicolour [Poaceae]) Landrace Storability in Ethiopia, *Economic Botany,* 53 (1): 69-78.

Thiele, G. (1999), Informal potato seed systems in the Andes: Why are they important and what should we do with them?, *World Development* , 27(1): 83-99.

Unruh , J. D., (2001), The Dilemma Of African Agrobiodiversity: Ethiopia And The Role Of Food Insecurity In Conservation, paper presented at the conference Managing Biodiversity in Agricultural Ecosystems 8-10 November 2001 Montreal, Canada.

US Bureau of the Census based on statistics from population censuses, vital statistics registration systems, or sample surveys pertaining to the recent past and on assumptions about future trends (2004)

USDA. Food Security Assessment, GFA 16, (2005), Shapouri, S., and S. Rosen, (eds.), Washington D.C., Economic Research Service,

Van Dusen, E., (2000) *In situ conservation of crop genetic resources in the Mexican Milpa system*, Ph.D. Dissertation. Davis, CA, USA: University of California-Davis.

Vavilov, N.I., (1956), The origin, variation, immunity, and breeding of cultivated plants. In: Selected writings of N. I. Vavilov, *Chronica Botanica*, 13: 1-16.

Weitzman, M. L., (1992), On Diversity, *Quarterly Journal of Economics,* 107(2): 363-405.

Widawsky, D. (1998), Pesticide Productivity, Host-Plant Resistance and Productivity in China, *Agricultural Economics,* 19(1-2): 203-17.

Widawsky, D. A., (1996), Rice Production, Yield Variability, and the War Against Pests: An Empirical Study of Pesticides, Host-Plant Resistance and Genetic Diversity in Eastern China, *Stanford University, Ph.D.*

Zegeye, T., Taye, G., Tanner, D., Verkuijl, H., Agidie, A., and Mwangi, W., (2001), *Adoption of Improved Bread Wheat Varieties and Inorganic Fertilizer by Small-scale Farmers in Yelmana Densa & Farta Districts of Northwestern Ethiopia,* Agricultural Research Organization (EARO) and International Maize and Wheat Improvement Center (CIMMYT), Mexico, D.F.: Ethiopia.

Zimmerer, K. S, (1996), *Changing fortunes: Biodiversity and peasant livelihood in the Peruvian Andes,* Berkeley: University of California Press.

Annex 1

Index	Concept	Construction	Explanation
Count	Richness	D=n or m S =	Number of cereal crops or crop varieties grown in community in the year; n is the number of varieties and m is the number of crops
Shannon	Evenness or equitability (Both richness and relative abundance)	$D=-\Sigma \alpha_i \ln \alpha_i$ $D \geq 0$	α_i= area share occupied by ith cereal crop or crop variety in community or by household in the year
Berger-Parker	Inverse dominance (relative abundance)	$D=1/\max(\alpha_i)$ $D \geq 1$	Max (α_i) is the maximum area share planted to any single crop or variety in community or by household in the year

Annex 2: Household Survey Data collected by Unit of Analysis and Analytical Purpose

Plot Level Only	**Use in Analysis**
Number of plots and area by plot	1) Measure of household assets in terms of natural capital to be used as exp. variable
Rental out of owned plot area	
Irrigation (yes no for entire plot)	
Soil quality (texture, color, slope, aspect, altitude)	2) Control for effects of environmental conditions on variety choice, HH income
Primary manager of the plot	
Distance to house	
Last use of plot before 2002	3) Control for security of tenure effects on variety choices, incomes
Rented in plots	
Irrigation (yes no for entire plot)	
Soil quality (texture, color, slope, aspect, altitude)	
Primary manager of the plot	
Distance to house	
Last use of plot before 2002	
Owner of plot	
Length of contract for owner	
Type of contract	
Amount paid for rent by plot for various rental arrangements	

Crop data by Plot	**Use in Analysis**
Crops planted on plot	Calculation of net returns to crop; Identification of what crops where on plots if needed for calculating area shares
Share of plot in crop	
Intercropping of crop (yes no)	
Planting date	
Expected/actual harvest	
Production season (belg or meher crop)	
Identity of person who obtained seed	Use in measuring seed system (e.g. female vs. male linked)
For Sorghum and Wheat: Variety data by Crop and Plot	
Variety planted in plot	Measure of ag. biodiversity at plot level; can be aggregated to HH level; Measure of seed system/check on responses on variety source asked in other section
Quantity of seed by variety	
Source of seed: amount used & price	
For Sorghum and Wheat: Crop by Plot	
Amount of family labor used for 9 cropping tasks (land prep, planting, weeding, input applications, irrigation, harvest, threshing, transport)	Calculating net returns to land, labor capital by crop but NOT variety Number of varieties as an explanatory variable and look at the impact on net returns to each crop for wheat and sorghum
Amount of hired labor used by plot, activity used for and cost	
Amount of exchange labor used for plot, activity used for and cost	
Working days of animal power used by plot, source of animal power and cost	
Use of other tools by plot, average cost	

For Sorghum and Wheat: Crop by Plot by Production Season	
Harvest date	Measures of crop failure and replanting Net returns to crop
Amount harvested	
Quantity and price sold	
Time sold at	
Who it was sold to	Analysis of gender division in seed system
Decision-maker on selling seeds/grain	
Manager of the money from sales	
Quantity used for home consumption (feeding livestock, eating, sharecropping, returned to HCS)	Non-monetary returns to crop
Use of straw and/fodder other than livestock?	
Main stresses on the plot	Indicator of presence of risk and management Indicator of food insecurity
Amount of harvest lost by plot	
Frequency over the last 5 years	
Coping mechanism	
Amount harvest stored	
Storage of seeds vs. grain	Measure of seed system Net returns to crops after storage
Main person responsible for storage	
Storage losses	
Causes of storage loss	
Use and cost of storage treatments	
For all other crops: Crop by Plot	
Input use, amount, price and source for fertilizer, pesticides family labor, hired labor, exchange labor, tractor, animal, tool rental	Net returns to ag. besides wheat and sorghum
By Input (seeds, fertilizers, pesticides)	
Purchase Source	Measure of seed system performance and market integration: transaction costs
Distance to house	
Sales outlet	
Distance to house	
By Variety Only – for Wheat and Sorghum Only	
Number of years planted (continuously?)	Measures of seed system performance; Determinants of seed demand Willingness to pay for attributes Supply source, reliability, transaction costs by seed source Measure of diversity and access Identification of blockages/constraints in supply system
How did they hear about it?	
Always able to get it	
Source to obtain first time	
Source at last planting	
How acquired if not retained	
How far away was source from house	
How was seed transported back to house	
Another source preferred?	
Reason why alternative not used.	
Who selects the seed	
When is it selected	
Who decides about planting varieties	
What physical characteristics are looked for	
Which characteristics are most preferred	

Which characteristics are least preferred	
When is the last time the seed was replaced	
Why was the seed replaced	
Which varieties would they like to plant and main reasons?	
Why didn't they plant those?	
How did they know about them	
Lost varieties	
Why can't replace variety	
Varieties they won't plant	
Varieties they stopped planting	
Household level data	
Members by relationship, gender, age, schooling	Measure of human capital as explanatory variable in HH welfare and ag. biodiversity
Migration out of HH and income from migrants	Measurement of household income/food security 1) to develop explanatory variable on HH welfare 2) to use for alternative measurements of income level (e.g. compare house construction with asset holdings with net income) 3) to use as exp. variable in explaining ag. biodiversity decisions
Off farm wage income by HH member	
Income from Non-ag. self employment	
Income from other sources	
Agricultural assets ownership	
Ownership of non-ag assets	
Characteristics of residence	
Livestock ownership amount and value	
Livestock products output and value	
Use and Access to Credit	
Participation in HCS activities	
Participation in other NGO activities	Measurement of social capital; direct evaluation of HCS activities
Ethnicity, religion and kinship membership	

Annex 3 : Team List

Team Leaders:

1. Dr. Lisanwork Nigatu
2. Dr. Negussie Dechassa Robi
3. Abduletif Ahmed
4. Yosef Assefa

Enumerators

5. Amare Getaneh
6. Amare Kebede
7. Bacha Daba
8. Bogale Lemi
9. Bulti Tesso
10. Fekadu Geta
11. Getu Mulatu
12. Jamal Yousuf
13. Jemal Abdulahi
14. Keder Bati
15. Melaku Zewdie
16. Negatu Bekele
17. Seifu Degaga
18. Seyoum Mengistu
19. Seyoum Tesso
20. Teklemariam Keneni
21. Tsegaye Gossa
22. Usmael Ahmed
23. Wendimsiamregne Mekasha
24. Wondwossen Seyoum
25. Yusuf Mussa

Survey coordinators

26. *Yohannes Amare*
27. Zenebe Tsegaw Wondimu
28. Gebrezain

Appendix 1: HH survey instrument round 1 and code sheet

Appendix 2: HH survey instrument round 2 and code sheet

Appendix 3: Survey management

Appendix 4: Recommendation to team leaders and list of most common mistakes

Appendix 5: Enumerator guide

Appendix 6: Community survey instrument

Appendix 7: Workshop results

ESA Technical/Field Reports

TECHINICAL PAPERS

The ESA Technical Papers are produced by the Agriculture and Economic Development Analysis Division (ESA) of the Economic and Social Department of the United Nations Food and Agriculture Organization (FAO). The series presents ESA's field and or technical reports. They are made available to the public through the Division's website. The analysis and conclusions are those of the authors and do not indicate concurrence by FAO.

ESA

The Agriculture and Economic Development Analysis Division (ESA) is FAO's focal point for economic research and policy analysis on issues relating to world food security and sustainable development. ESA contributes to the generation of knowledge and evolution of scientific thought on hunger and poverty alleviation through its economic studies publications which include this working paper series as well as periodic and occasional publications.

Agricultural and Development Economics Division (ESA)
The Food and Agriculture Organization
Viale delle Terme di Caracalla
00100 Rome
Italy

Contact:
Office of the Director
Telephone: +39 06 57054368
Facsimile: + 39 06 57055522
Website: www.fao.org/es/esa
e-mail: ESA@fao.org