Using markets to promote the sustainable utilization of crop genetic resources:

Project Methodology

with contributions from:

Leslie Lipper, Leigh Anderson, Melinda Smale, Jon Hellin, Toby Hodgkin, Tim Dalton, Conny Almekinders, Patrick Audi, Mauricio Bellon, Romina Cavatassi, Lamissa Diakete, Richard Jones, Madelon Meijer, Latha Nagarajan, Alvaro Paz, Monika Rodriguez, Amadou Sidebé, Joost van Heerwarden, Paul Winters

Agricultural and Development Economics Division
The Food and Agriculture Organization
of the United Nations
http://www.fao.org/es/esa/

A Project Funded by the FAO-Netherlands Partnership Program

Using markets to promote the sustainable utilization of crop genetic resources:

Project Methodology

Abstract

This document is the result of the combined efforts of inter-disciplinary FAO-CGIAR partners. Many of the foundations for this document were developed in a series of three project workshops held in 2004, 2005 and 2006 at FAO headquarters. This document aims to provide some guidance on the design and the implementation of the methodology conducted to evaluate the relationship between agricultural markets and farmer's decision making. Specifically, it provides a summary of the conceptual and empirical frameworks that have been adopted in order to empirically assess the impact of markets on farmers utilization of crop genetic resources, and to identify possible interventions that could be implemented by the public sector to promote the sustainable utilization of crop genetic resources. Different concepts and instruments such as value chain analysis, sampling strategies, vendors, community and household data collection, among others, are carefully described and analyzed. In addition, the different econometric techniques that could be considered by researchers when assessing the impact of on-market biodiversity and on-farm biodiversity on farmers' welfare are also examined. The concepts, the methodology, the instruments and the econometric approaches described below are not exclusive for this particular research and can be implemented in other settings.

Key Words: Crop Genetic Resources, Biodiversity, Markets, Value Chain Analysis, Diversity Measures, Welfare Measures.

Foreword

This methodology document combines the contributions of many individuals who participated in project workshops and who offered written contributions. Any misattribution or lack of attribution, of any one person's contribution is wholly accidental as this document has evolved over time.

This document is divided into major sections. The first section describes the context and issues motivating this project while section II describes the empirical methodology used in the value chain, market and household data collection activities.

Table of Contents

LIST OF TABLES:

List of Boxes:

List of Figures:

I. Objectives, concepts and definitions, and research questions

A. Introduction

In 2004 the Agricultural and Development Economics Divisions (ESA) of the FAO initiated a research program on agricultural markets and their relationship to farm level decisions on utilizing crop genetic resources. The program is motivated by the need of policy-makers in developing countries to respond to commitments made under the International Treaty on Plant Genetic Resources to promote the sustainable utilization of plant genetic resources. This research program is designed to support and inform strategies to meet this commitment.

The program builds upon considerable work that has already been done by FAO, CGIAR centers and others on seed systems, agricultural biodiversity and farmer livelihoods. Several CGIAR centers, including IPGRI, IFPRI, CIMMYT, CIP, and ICRISAT are partners in the program, and actively involved in developing the project methodology and implementing case studies in the field. This document is the result of an inter-disciplinary collaborative effort amongst project partners. Many of the foundations for this document were developed in a series of three project workshops held in 2004, 2005 and 2006. These have now come to fruition in this document and in the implementation of the methodology in the field.

Writing this methodology was an interative process. It took place during the case study implementation, and we have been able to incorporate the experiences of grappling with field reality in writing up the methodology. This has resulted in a very rich and practical guide.

B. Project Objectives and Context

The objective of this project is to empirically assess the ways in which agricultural markets affect how farmers utilize crop genetic resources, and to identify possible interventions the public sector could take to improve the market-stimulated farmer incentives to use these resources in a sustainable way. Ultimately the goal of the project is to identify ways to promote the sustainable utilization of crop genetic resources, which is a primary objective of the International Treaty on Plant Genetic Resources for Food and Agriculture and the 111 countries which are parties to it.

Participating in agricultural markets is one of the most important determinants of the way farmers utilize crop genetic resources: e.g. the crops and varieties they choose for planting. Participation in agricultural markets also has significant impacts on farm welfare. Markets are becoming increasingly important in the agricultural sector of developing countries and hence understand their role in the management of crop genetic resources is fundamental in sustainable utilization. In the last two decades, the implementation of structural adjustment reforms and liberalization has greatly increased the influence of market institutions on farmers. At this point, markets are seen as the primary engine for promoting agricultural development, with attention now turning to making markets work for small and low income farmers.

It is important to recognize that markets are governed by a set of formal and informal rules and relationships, and these are affected by public policies and specific interventions. It is also important to recognize that one of the most pressing current debates in development is how best to govern markets to achieve economic growth and development. One major concern of economists has been improving the efficiency of markets by reducing transaction costs and providing a supportive institutional environment. (Jayne et. al. 2001; World Bank 2002) The role of transaction costs in a smallholder farmer's decision to participate in input

and product markets has also received considerable attention (Key, Sadoulet and deJanvry, 2000; Sadoulet and deJanvry, 2003; Bellemare and Barrett, 2006). In particular, attention has focused on the identification of behavioral responses to adverse conditions characteristic of many markets in less developed countries and in policy options to reduce transaction costs and improve the functioning of markets. Others have argued that non-market mechanisms may be needed, at least initially, to overcome the risks facing market participants in poor rural areas, particularly in the case of staple food crops (Dorward, Kydd, Morrison & Poulton 2005) While there is no consensus on how best to govern agricultural markets in developing countries, there is clearly agreement that markets are not working well in many contexts and interventions are needed to improve their performance (Dorward et al., 2005)

In this project we are building upon the work on how to govern agricultural markets for agricultural development. However we are broadening it by adding crop genetic diversity as an important indicator of market performance - as an indicator of the range of choices available to farmers in using markets as a seed source. The addition is not trivial as markets are not only a vehicle for exchanging seeds – a physical good – but also for the exchange of the crop genetic resources embodied in seeds and expressed as varieties. Thus we need to consider how well markets work as a distributor and point of access for both seeds and genetic resources. We will be assessing seed market structure, conduct and performance and relating this to farm level use of seeds and varieties and the impact of farm use patterns on farm welfare and on-farm diversity. One main innovation in the study is to measure crop genetic diversity in the market chain and consider it as an important determinant of market performance. Another is considering the impact of market exchanges on both private and public goods: e.g. the impact of increased access to crop genetic resources on: 1) the farm household's crop productivity, resilience and consumption patterns and 2) on-farm crop biodiversity conservation. Both of these factors are important components of a sustainable pattern of utilization.

We can think of market exchanges between farmers and seed suppliers as lying on a spectrum from formal to informal, defined as the degree to which the rules of exchange are codified and/or standardized and thus, applicable and known across a broader population of participants. An informal exchange may be governed by a set of cultural norms or ad hoc rules that are determined and known only to the participants to the exchange. A formal exchange may be subject to national and municipal regulations, therefore having terms that are not completely determined by the participants.

The market for seeds is different from many product markets because characteristics, embodied with the seeds, are difficult to identify at the time of purchase. Production or end-use attributes are not readily observable, nor often explicitly priced, and revealed piecewise to the purchaser after the seed has been planted, grown and harvested. Seeds are not "transparent," making the consumption and production attributes of seed that farmers value, its quality and genetic content, difficult to measure. (Morris, Rusike and Smale, 1998,) The degree to which any of these features are observable to farmers varies by crop and variety, but lack of transparency is a problem that results in information asymmetries between the consumer (farmer) and supplier of the seed (Morris, Rusike and Smale, 1998). Exchange will thus be tempered to the degree that non transparency renders property rights incomplete and the transaction costs of measuring attributes and ensuring the value of the seed, positive.

Research has indicated that the existence of such information asymmetries is a major factor in farmers' decision to exchange in the informal sector, where the expression of genetic content in specific environments can be observed, and where social relations and reputation provide

some assurance of quality (Badstue, Bellon, Juárez, Manuel, and Solano 2004). It has also been shown as a source of inefficiency between farmers and end-users (Lambert and Wilson, 2003; Barkely and Porter 1996).

To farmers, the decision to participate in the seed marketplace and purchase and grow a particular seed (and the embodied attributes of the crop or variety) has a significant impact upon household income, consumption and household production activities. Because of the limited observability of the public and private genetic benefits embodied in seeds, the potential for an Akerlovian[1] lemons market exists and might create a situation where the "bad" may drive out the "good."

The "lemons" market for seeds has been managed through the existence of parallel marketing channels. The first marketplace is the formal certified seed market (for example the hybrid seed marketplace) while the second is the informal purchase or exchange market. The informal seed market is governed by social networks and relationships where the reputation of the seller substitutes for the certification, seal or guarantee of the formal marketplace. Generalizations are not easy to make since the choice, and predominant marketing vehicle is often crop-, location- and time- (in a historical sense) specific (see McCann (2005), for example, on the rise and fall of the hybrid maize market in Zimbabwe). More recently, and owing to the failures in the formal marketplace, an intermediate market channel, characterized by less rigorous quality control (and hence reputation) but more local production responsibility (and hence greater social network recognition) has evolved to emerged under certain conditions.

In the Akerlovian lemons market, asymmetric information about the true characteristics of a good leads to the situation where poor quality goods largely displace high quality goods. In the seed marketplace, poor quality could be manifest in a private dimension, for example in germination rates, "cleanliness," or local adaptation, and also in a social dimension. The social dimension might be thought of as the value of the genetic attributes that make varieties differ in their productivity, resiliency, or uniqueness. The Irish Famine of the 1840's, where nearly all cropland was planted to two potato varieties, Lumper and Cup, is an example where crop diversity narrowed to an extreme state and where the lack of disease resistance caused devastating human losses (Ó' Gráda, 1999).

The research methodology described in this paper is a first attempt to grapple with these issues. The project is being implemented through five pilot case studies where empirical information is being gathered on the relationships that exist between agricultural markets, their governance and the nature of the exchanges that take place within them, and then to assess how market characteristics condition farmers' access to crop genetic resources in markets and ultimately the crops and varieties they plant in their fields, e.g. their utilization of crop genetic resources. As a final step in the methodology, on-farm utilization patterns are being linked to different dimensions of household welfare and crop diversity in order to better understand both the private and social impacts of agricultural seed market performance.

C. Concepts and definitions

[1] George Akerlof won the Nobel Prize for Economics in 2001 for his 1970 essay "The Market for Lemons" which provides a formal analysis of the impact of information asymmetries on market efficiency. When sellers have more information about the quality of a good than the buyer, the entire market may collapse or may contract into an adverse selection of low-quality products – e.g. a markets for "lemons".

Much of the methodological work and discussions held in workshops and electronically among project participants has focused on refining a set of key concepts and definitions: markets, welfare, crop genetic diversity and access to crop genetic resources.

1. Markets

The study is focusing on markets where seed is exchanged between a buyer and seller in a voluntary transaction and they may involve seeds from either the formal or informal sector (e.g. certified seed of improved varieties, or local landrace varieties) We are focusing on retail transactions – e.g. the farmer is the purchaser of the seed. The definition of the market will vary in each case study but basically we will be looking at the "top" of the market chain : e.g. retail or "micro-retail"(see Fafchamps and Vargas, 2005). We are interested in how variations in market characteristics affect access to CGR, and how those market variations are, or are not, susceptible to policy. The market variations of primary interest are scale and composition (which may be affected by general agricultural development policies or regulations on market integration, competition or infrastructure, or specific seed sector policies), and the presence of direct seed sector interventions such as seed fairs or emergency seed. The markets we want to concentrate on in this study are exchange nodes that have a role as a hub in dispersing CGR. We are interested in how buyer and seller preferences, coupled with variations in policy and physical constraints and opportunities, shape the market and access to CGR – availability, cost, and information. Specific market characteristics of interest therefore include the dispersion of CGR, market scale (size and frequency), market composition (of participants and seeds from either the formal or informal sector), and the institutional and physical environment in which the exchange takes place.

Examples of location and characteristics of vendors in retail markets sites in various case study sites:

Chiapas (seed input shops, despachos, veterinarians)
Kenya (stockists, traveling traders, farmers selling surplus in open market)
Mali (weekly village market)
India (shandies, village markets)
Bolivia (spot markets in villages; larger permanent markets)

2. Welfare

Welfare will primarily be defined at the farm household level (e.g. this will be the project standard) although information on intra-household and community level measures of welfare could optionally be gathered in some of the case studies. The definition of farmer welfare used in the study is broad and linked to literature on the value of maintaining diversity on-farm (Smale 2006). We are using a broad definition of welfare to include not only asset and income based wealth measures, but also indicators such as resilience to production and consumption shocks, and dietary diversity.

3. Access to crop genetic resources in markets

This concept is key (and unique) to our project. We define access as having 3 dimensions: physical availability, information and cost. Crop genetic diversity is our primary measure of physical availability. We define crop genetic diversity (in general) as a critical dimension of farmer access to crop genetic resources and we are interested in measuring the crop genetic diversity in one important piece of the seed system (markets) in order to better understand farmer access to crop genetic resources from markets. We are going to try to measure access to crop genetic resources in markets by collecting information on the genetic diversity in the

market over a specific time period, as well as information available about that diversity and the costs of obtaining it.

4. The sustainable utilization of crop genetic resources

In this study we consider the sustainable utilization of crop genetic resources to be a pattern of utilization of crops and varieties that generates three main categories of benefits:

- Private benefits to farmers via the consumption and production "services" that genetic resources (in interaction with human and environmental factors) provide to farmers
- Global public benefits in the reduction of genetic erosion (maintaining genetic option values) and
- Local public benefits of reducing genetic vulnerability (increasing resilience).

The scale at which these benefits are realized vary. In the first we are looking at farm level benefits, for the second the relevant unit is a genetic population over time and for the third the relevant scale is regional. There are likely to be tradeoffs between the three categories of benefits and assessing these is one of the intended outcomes of the study.

D. Project research questions

The project methodology is being designed to answer three main research questions via case studies. These have been developed through a consultative process and discussed in the three project workshops held in 2003-2006.

1. **How do we measure farmer access to crop genetic resources, especially in relation to crop genetic diversity, in local markets?**

2. **How does access to CGR in markets affect farmers' decisions to participate in markets for seeds, and what impact does this have on farm levels of welfare and on-farm crop biodiversity?**

3. **How do policies and seed system interventions such as emergency seed relief, seed sector regulation and seed and diversity fairs affect farmer access to CGR in local markets?**

The methodology which we are proposing to answer these questions and which is being tested through the implementation of the pilot case studies is the subject of the second part of this document.

II. Methodology

A. Components and Sequencing

To answer the research questions outlined in section I, we are developing a project methodology for a set of case studies. The methodology has 3 components: 1) identifying and quantifying the access to diversity in a set of retail markets, 2) assessing the socio-economic, legal and regulatory factors that are likely to influence the market level of access to diversity and 3) relating the market level of access to diversity to farm level utilization of crop genetic resources and measures of on-farm diversity and welfare.

Component 1: Identifying and quantifying the access to diversity in retail markets

Task: Identifying a set of retail seed outlets for seeds, assessing the availability of crop genetic resources in the markets for a specific period of time, and assessing the information provided about the resources, as well as costs of obtaining them in the market. Collecting information on market structure, conduct and performance.

Method: Value chain analysis, market observation, key informant and vendor survey, focus groups with vendors

Component 2: Assessing the socio-economic, legal and regulatory factors affecting access to crop genetic resources in markets

Task: Analysis of the input and output value chains for the selected crops together with an analysis of the legal and regulatory environment governing plant varieties, seeds and markets.

Method: Value chain analysis using key informant and focus group techniques (with some overlap on the key informant survey in component (1)),
Policy and regulatory analysis using key informant analysis, published information

Component 3: Relating the market level of access to diversity to farm level utilization of crop genetic resources and measures of on farm diversity and welfare.

Task: Collect farm household level data on diversity and welfare from households located within the market zone.

Method: Household survey, farmer focus groups in communities, quantitative analysis of household outcomes using variables created from market level data

Table 1, summarizes the components and sequencing of the methodology – based on discussion held in the March 2006 project workshop.

Table 1 : Steps in implementing the project methodology

Activity	Remarks
1. Site/target crop selection	Using project criteria: Study sites and target crops should be located in areas where the crop grown is important for food security, has significant levels of local genetic diversity and for which markets are a significant source of seeds.
2. Value chain Analysis (VCA)	First research activity to be implemented as it yields important information needed to design the subsequent analysis. The objective is to map both input and output market channels for the target crop and to provide a basis for understanding how crop genetic diversity is being transmitted through the market chain and why. Results will be used to select the sample of retail market sites and vendors for market surveys (3). The VCA should also highlight key policy, regulatory and other types of blockages in the system that should be further investigated in the policy and regulatory analysis (5). The method involves the use of both qualitative and quantitative data collection methods during focus group meetings and key informant interviews. Farmer focus groups should also be used to identify the entire range of seed sourcing channels (e.g. both market and non-market), their strengths and weaknesses as well as to describe the markets. The VCA can also be used to provide temporal and historical perspectives which are an important complement to the cross-sectional quantitative work. Specific information to be collected: • Identify various market participants and their functions. • Identify channels in seed and grain markets at project sites going from consumer (e.g. farmer) to source (e.g. local, national, international). • Indicate volumes, quantities, grades, varieties, attributes and marketing margins for seed at various points in the chain. • Include formal and informal seed sector
3. Seed and grain retail market sample surveys	Data to be collected on market structure, performance and conduct, including measures of crop genetic diversity. Three instruments have been developed for data collection: the market observation, key informant and vendor surveys. The key informant survey may be thought of as a part of the VCA using a structured interview protocol. Timing for collecting these data is critical. The vendor survey must be done when seed is actually being sold on the market and this period needs to be determined via the value chain and focus group or pre-survey. The surveys are intended to gather information on market participants,

	crops, varieties, attributes, seasonality, quantities, prices for the retail market outlets for seeds. The data collected in these surveys will be used to develop market measures of diversity, information and prices for quantitative analysis. Should include focus groups of farmers purchasing in the market to get info on traits desired, transactions costs they face, as well as focus groups of vendors to collect information on their perceptions of traits associated with different crop varieties.
4.Household survey	Should be preceeded by farmer focus groups establishing local distribution of diversity (not necessary for sites where local diversity already characterized); also key factors determining market participation. It should be done immediately after the seed market survey or after harvest. Sample should include both participants and non-participants in the market for seed Data to be collected on market participation, welfare, on farm diversity and general household demographics. Data to be used together with data from market surveys to analyze the impact of access to crop genetic resources in markets on farm level participation in markets and implications for on farm patterns of crop utilization, welfare and diversity.
5.Policy and regulatory analysis	Could be done at any point after the VCA which should highlight the key issues to be focused upon in the analysis. Should cover not only seed regulation but also broader policies affecting the supply and demand of agricultural products and inputs and the development of agricultural markets.

B. Value Chain Analysis

1. Introduction

Value chain analysis (VCA) is one of the key methodological tools the project is using to answer questions about the relationships between diversity in markets and farm level access and utilization of crop genetic resources. Conducting the value chain analysis is the first step in the project methodology and it provides much of the basic information needed to move to subsequent stages of collecting quantitative data from households and markets and an in depth review of the policy and regulatory environment. The value chain analysis is used to map out relations and flows of crop genetic resources (e.g. varieties) and seeds in formal and informal markets and how they relate to flows in output markets. The key questions addressed by the VCA are::

- Which seed varieties are available to farmers through formal (and informal) markets?
- Whether the seed varieties supplied genuinely reflect farmers' needs?
- To what extent do output markets for crops influence the varieties that farmers grow (and, hence, the types of seed that they require)?

The first step in answering these questions is to 'map the market', to build up an understanding of the different players or actors in the seed input and product (e.g. grain or tubers) output chains and the relationships between them, along with the factors that determine how well or poorly the chains are working. We need this understanding in order to shed light on some of the factors determining why farmers are purchasing different types of seed. An understanding of the different actors also helps us identify where some of the other research tools, such as the vendors' survey, should be directed.

These guidelines are based on a qualitative approach to mapping value chains that is being used in Mexico, Bolivia and Ecuador. The value chain analysis is an iterative process and while predetermined topics for discussion can be identified, it is far harder (and somewhat restrictive) to try and prescribe specific questions that those working on other case studies should use. What is provided here is a conceptual framework – the market map –and some insight from experiences in the field in Mexico, Bolivia and Ecuador.

2. What is a value chain?

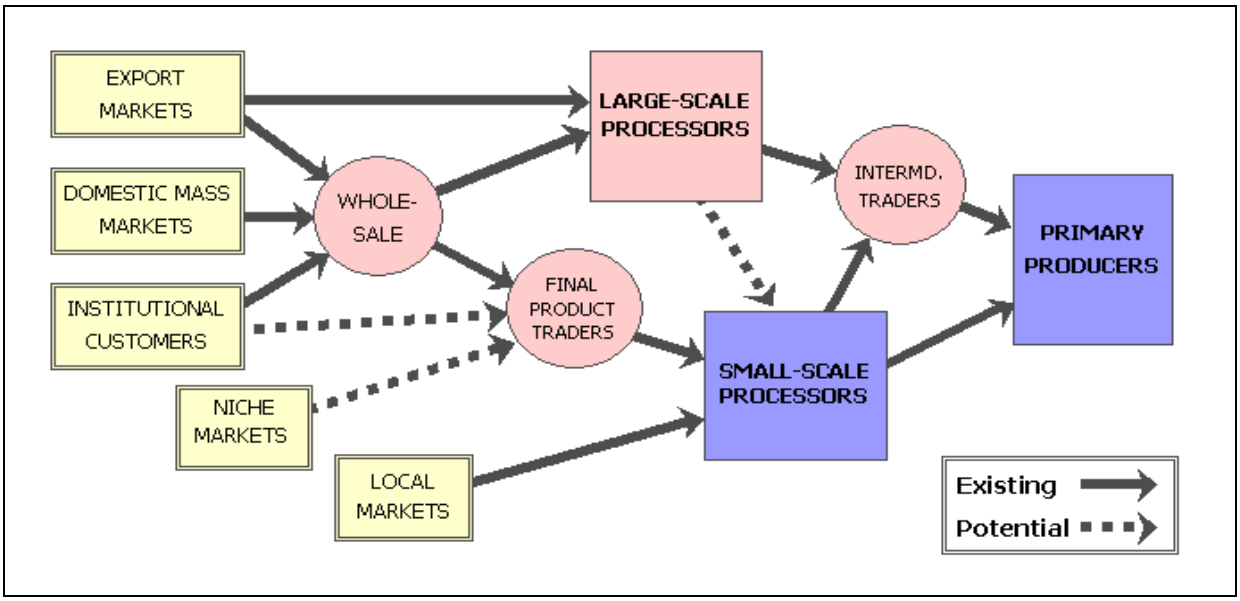
The flow of seed to farmers (e.g. the input market) and grain or tubers from farmers to the market (e.g. the output market) occurs along chains. These can be referred to as value chains because as the product moves from chain actor to chain actor e.g. from producer to intermediary to consumer, it gains value. A value chain can be defined as the full range of activities which are required to bring a product or service from conception, through the different phases of production (involving a combination of physical transformation and the input of various producer services), delivery to final customers, and final disposal after use. The chain of actors who actually transact a particular product as it moves through the value chain includes input (e.g. seed suppliers), farmers, traders, processors, transporters, wholesalers, retailers and final consumers. A simplified version of a value chain is shown in Figure 1.

Figure 1 A simplified value chain

Seed suppliers → Farmers → Traders → Processors → Exporters/importers → Retailers → Consumers

In reality, value chains are more complex than the above example, in many cases, the input and output chains comprise more than one channel and these channels can also supply more than one final market. In some cases input and output chains are linked – in others they are totally separate. A comprehensive mapping therefore describes interacting and *competing* channels (including those that perhaps do not involve smallholder farmers at all) and the variety of final markets into which these connect (see Figure 2).

Figure 2 A more complex value chain of the output market



*Defying convention, Figure 2 reverses the direction of the chain. It shows the flow of *income* from markets along the chain to primary producers, rather than (as is conventional) the flow of *goods* in the opposite direction. This counter-intuitivism emphasizes a demand-led perspective.

3. Research tools for conducting VCA

Value chains can be mapped and analyzed using value chain analysis (VCA) which can include qualitative and/or quantitative tools. There are no fixed rules on which research approach is better but there are strong grounds for recommending that a qualitative approach is used first, followed (time and resources permitting) with a quantitative study (see Box 1). The analogy is one of painting a house: the first coat (the undercoat) is provided by short qualitative study (guidelines for the qualitative research *per se* are given below).

The initial study adds a little color but several coats of paint are needed in order to appreciate the final effect. What we have done to date is the equivalent of our undercoat. We can see who the different value chain actors are but we we do not know the relationships between them, the prices and quantities of seed moving through the first bit of the chain, the crop diversity found in farmers' fields, the prices and quantities of grain or tubers as they move to

the right of the chain, the rationale for why farmers are purchasing seed *x*, *y* or z. We need more color; we need more layers of paint. We can add this color via qualitative (e.g. semi-structured interviews and focus groups) and/or quantitative (e.g. household survey or a questionnaire) tools (see Figure 3). If time and funds are short it may be best to focus on qualitative research bearing in mind that a great deal of information on prices and quantities can still be gleaned from qualitative research and often secondary sources such as national statistics.

Box 1: Unlocking the complexity of value chain actors' realities: combining qualitative and quantitative research

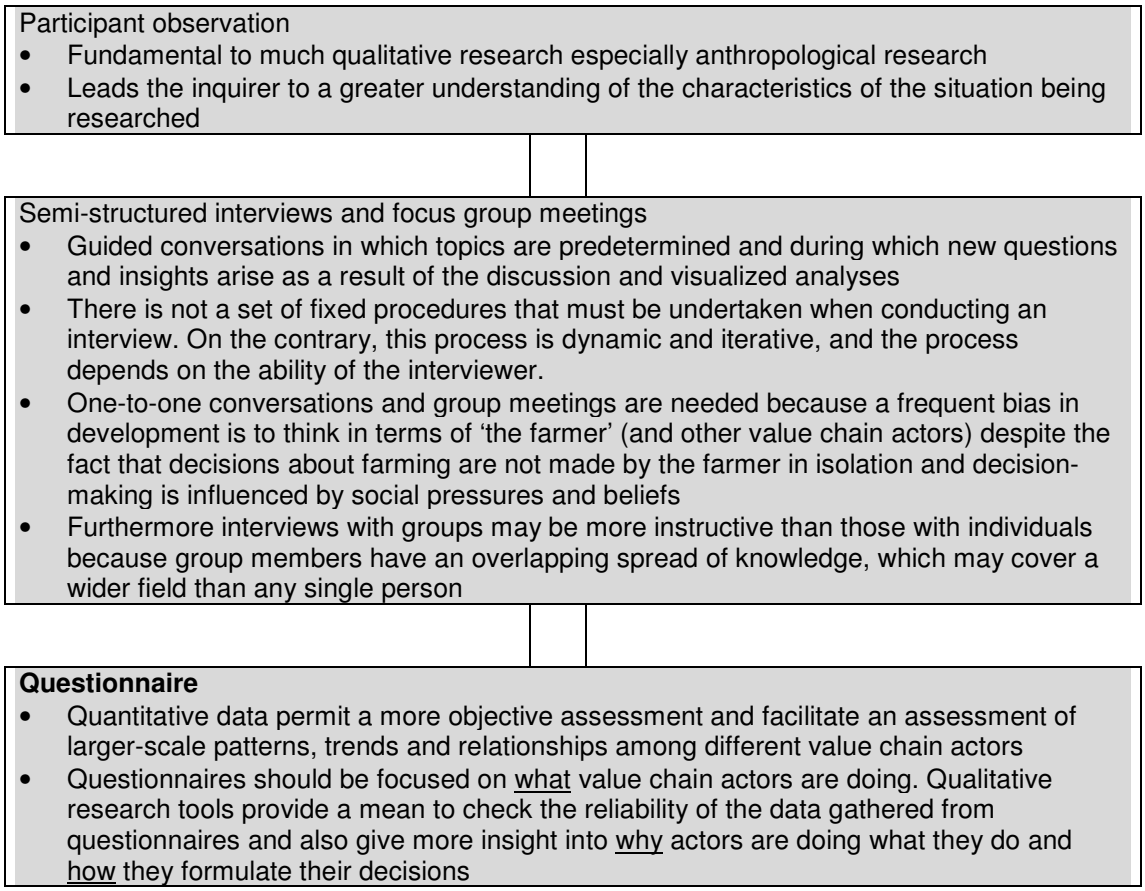
Endeavours to understand the realities of different value chain actors are plagued by difficulties. Inevitably, researchers are dependent on information from the different actors themselves. We need to question continuously whether the indicators developed are valid *i.e.* do they measure the concept they are designed to measure, and whether the information we collect is reliable *i.e.* a question is of little use if some of the value chain actors, such as farmers, answer it in one way one day and another the next. Ensuring a high degree of validity and reliability is one of the persistent concerns in any social science research strategy. It can be particularly difficult in the context of smallholder agriculture and value chains.

One of the most widely used quantitative research tools are questionnaires. There are many advantages to questionnaires of course but the drawback of relying exclusively on a research tool such as a questionnaire is that there is no way in which increased rigour during analysis can compensate for the unknown and degree of inaccuracy involved in the measurement process. Furthermore, questionnaires may entail interpersonal relationships of power and distort value chain actors' realities by fitting them into centrally pre-set frameworks. Questionnaires may also suffer from the same degree of subjectivity as that normally attributed to qualitative research by reflecting the predisposition of the researcher.

Questionnaires often fail to capture many of the nuances of actors' realities, the reason being that their knowledge systems are often not verbally or numerically codified. In this case, qualitative data, such as that gained by participant observation techniques and semi-structured interviews may better represent their perceptions and realities. Indeed, observation, interview and casual conversation also cause less suspicion and less guarded comment than research methods that involve outsiders writing down responses.

In this project we have adopted an approach that combines quantitative and qualitative research tools (see Figure 3). When quantitative studies are combined with a credible understanding of complex real-world situations that characterise good qualitative studies, we can gain a sound understanding of the problems and opportunities faced by different players in the various value chains that we are focusing on. A real understanding of the way that a particular value chain works can 'unlock doors': farmers, processors and other value chain actors may well judge us on the basis of our behaviour, attitudes and questions, hence, irrelevant and culturally insensitive questions can result in scepticism, distrust and lack of co-operation. Much can be gained by building up an understanding of how these chains work before designing a questionnaire.

Figure 3. Tools for value chain research



4. The Market Map

If we want to understand more about the rationale behind farmers' decisions *vis-à-vis* the types of seeds that farmers purchase etc. then we also need to know about the extraneous factors that influence the way that the value chain works. This is where the market map is useful. The market map is a conceptual and practical tool that helps us identify policy issues that may be hindering or enhancing the functioning of the chain and also the institutions and organizations providing the services (e.g. market information, quality standards) that the different chain actors need in order to make better informed decisions.

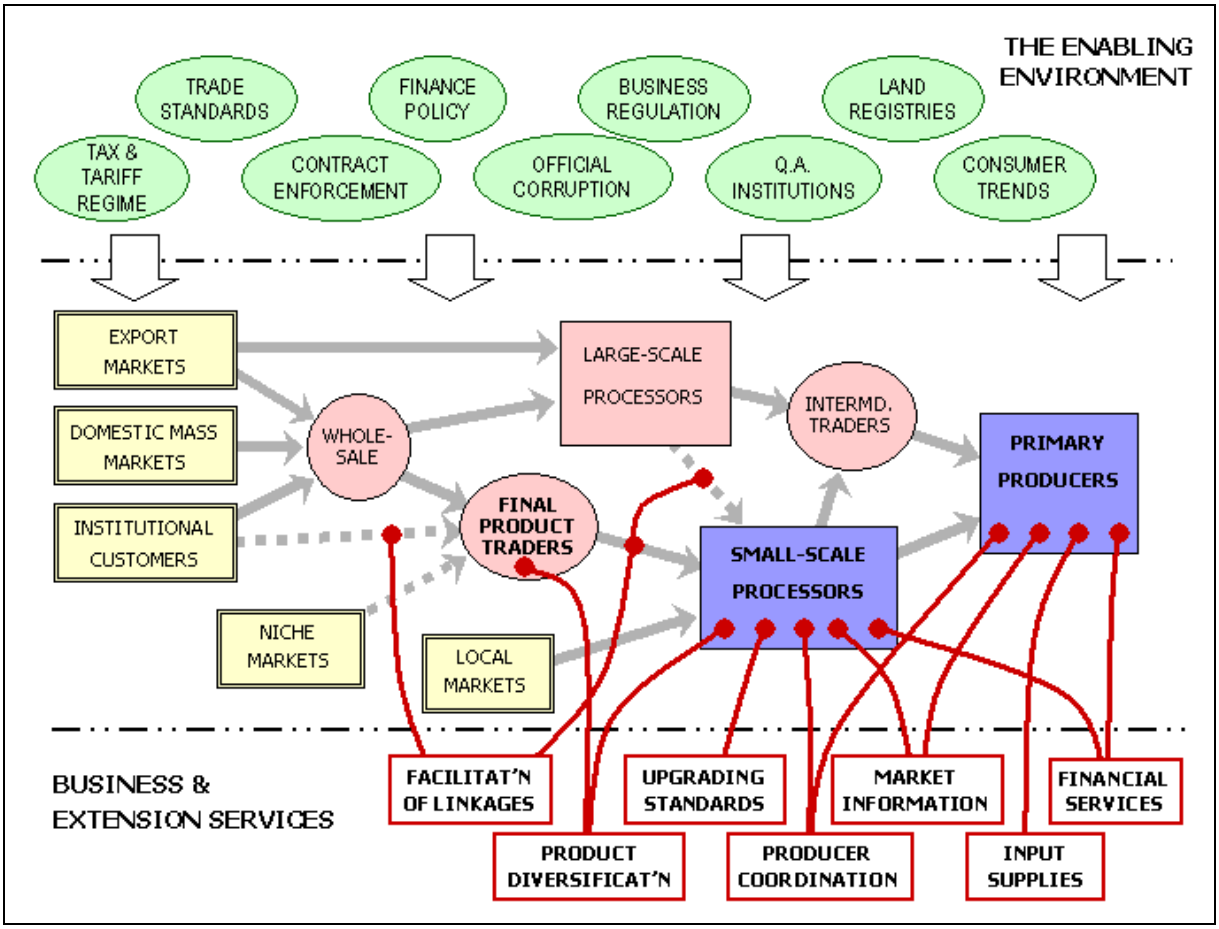
For example, a group of farmers may not know that a particular seed supplier has on offer a seed type that no other seed supplier has in stock. If the farmers do not know the seed is on offer, they may not buy it and, consequently, that particular variety will not be planted. Another example is that farmers might hear from the radio that there is an increasing demand for a particular type of maize. On hearing this on the radio they may well then go and seek out seed of the maize type in question. In order to understand farmer decision-making *vis-à-vis* what seed they purchase, it is important to identify the farmers' sources of information.

The Market Map is made up of three inter-linked components (see Figure 4):

- Value chain actors (see above)
- Enabling environment (infrastructure , policies, institutions and processes that shape the market environment)
- Service providers (the business or extension services that support the value chains' operations)

The enabling environment consists of the critical factors and trends that are shaping the value chain environment and operating conditions, but may be amenable to change. These "enabling environment" factors are generated by structures (national and local authorities, research agencies etc.), and institutions (policies, regulations and practices) that are beyond the direct control of economic actors in the value chain. The purpose of charting this enabling environment is not simply to map the *status quo,* but to understand the trends that affect the entire value chain, and examine the powers and interests that are driving change. This knowledge can help determine avenues and opportunities for realistic action, lobbying and policy entrepreneurship (admittedly activities that we are unlikely to be engaged in as part of this project).

Figure 4 The full market map



In the most effective value chains, the actors who actually form the chain (i.e. transact the main product) are supported by business and extension services from other enterprises and support organisations (e.g. seed suppliers and intermediaries). There is an on-going need for chain actors to access services of different types both market and technical. The third component of the Market Map framework is concerned with mapping these services that support, or could potentially support, the value chain's overall efficiency. The services that can potentially add value are many and include:

- Input supplies (seeds, livestock, fertilizers etc.)
- Market information (prices, trends, buyers, suppliers)
- Financial services (such as credit, savings or insurance)
- Transport services
- Quality assurance - monitoring and accreditation
- Support for product development and diversification

We have already introduced above the middle layer of the market map – the value chain – but the other layers – the enabling environment and service providers – have a big impact on how the value chains function and, hence, in our case the impact of markets on crop diversity (see Box 2).

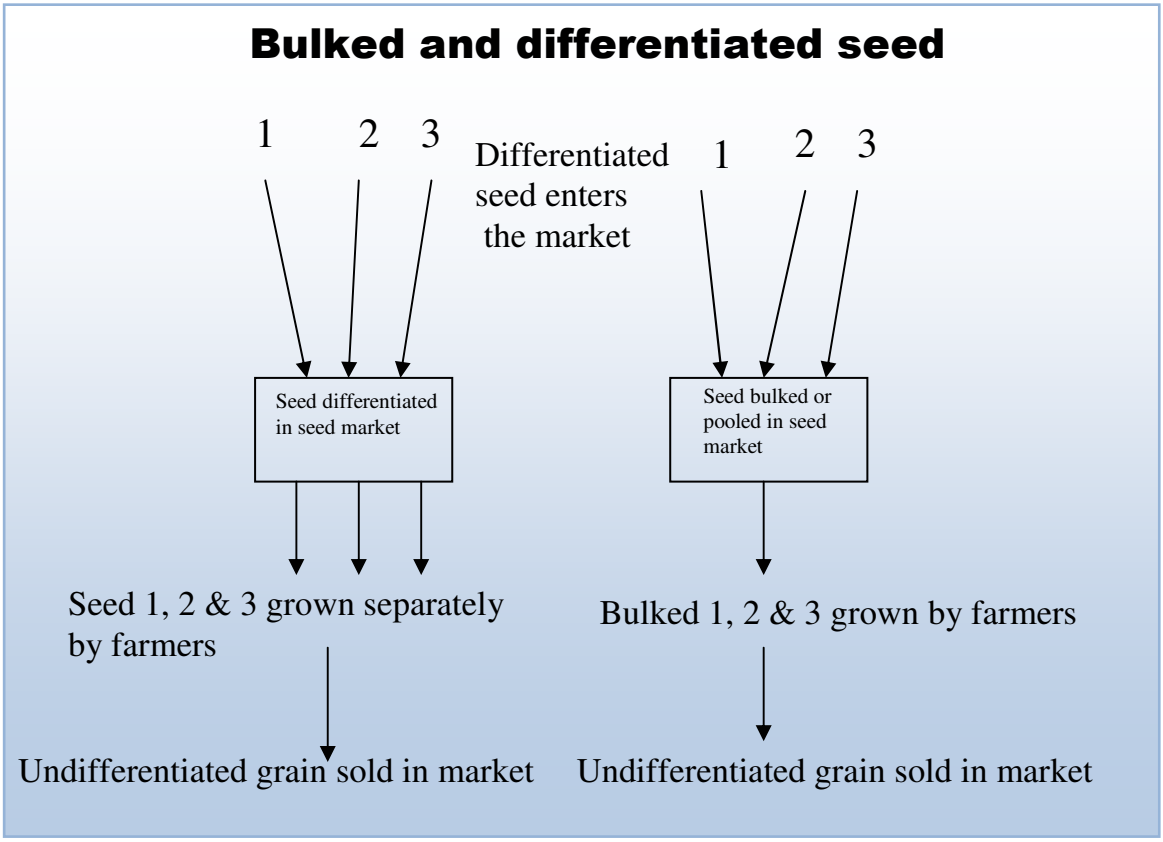
Box 2
The enabling environment, service providers & crop diversity in Mexico

The market map, proved to be a useful tool for the maize VCA in Mexico. Farmers and seed suppliers pointed out to us that seed availability is very much influenced by the subsidies that the state and federal government make available to the producers of hybrid and open pollinated varieties (OPV) of seed. There is evidence that these subsidies are undermining farmers' traditional seed recycling practices: with the subsidy, OPV seed is free so there is little incentive to continue growing land races when you can get hold of a 20 kg bag of treated seed at no cost. In the context of the market map, one of the biggest influences on what seed farmers purchase is the policy environment.

5. Identifying bottlenecks

Value chains do not always work efficiently and effectively (either in terms of income generation for small-holder producers and/or the maintenance of crop diversity). The VCA can help researchers identify where there are bottlenecks. At the project workshop a hypothetical example of two different situations was discussed (see Figure 1). The first example where the seed market maintains differentiated seed is based on the actual case in Chiapas. Figure 1 could be quickly built up using qualitative research tools such as semi-structured interviews or even very casual conversations with different groups e.g. farmers, seed companies and grain purchasers.

Figure 4. Two examples of seed and grain value chains



The chain on the left of Figure 1 was mapped in one day. Figure 1 can, therefore, be seen as the 'undercoat'. It may be the case that we would be happy with this amount of detail but it is likely that we would want to know more about why the chains are as they are. Bearing in mind the main theme of the research project (agricultural markets and their relationship to farm level decisions on utilizing crop genetic resources) we would probably want to know the following:

- Why do farmers in the left hand chain sow seed lots separately when the grain they sell is pooled by the grain merchants? Here is where theorizing is helpful in guiding instrument design. If, for example, we think that part of the answer lies with the regularity with which they are net sellers (i.e. for those who sell infrequently it doesn't matter), then we make sure that we have a question on how often they are net buyers or sellers.

With respect to the right hand chain

- Why is the seed pooled in the market? Do traders mix up varieties because they do not know what differences are between them? Do traders mix seed because it is cheaper for them to deal with the seed this way? Are there extraneous policy issues that account for why traders bulk seed i.e. a government subsidy for the number of bags of seed sold irrespective of what type of seed it is?
- Would farmers prefer the seed to remain differentiated?
- Do traders understand that farmers would prefer that seed lots be kept separate?
- Would farmers be prepared to pay more for differentiated as opposed to pooled seed?
- Do traders know that farmers would be prepared to pay more for differentiated seed?

- How much more would farmers have to pay for differentiated seed in order to make it worthwhile for traders to provide this?
- Would farmers be prepared to pay more for differentiated seed if there was a market for differentiated grain?
- Would traders be able to sell more seed (and at a higher price) if they provided farmers with differentiated seed and information on the seed? i.e. are the same quantities of seed per farmer in the left hand and right hand chains?

The above is an example of the types of questions that should be asked in the value chain study. Answers to these questions can be gleaned by using the aforementioned qualitative and quantitative research tools. In this case, we are adding more coats of paint and by doing so we are seeing more details (more color). By adding more color, we can document the transaction costs involved in selling seed, the prices and quantities of seed moving through the chain and the crop diversity found in farmers' fields etc.

C. Analysis of the Policy and Regulatory Framework

This analysis can be done at any point in the project implementation after the value chain analysis has been conducted. The purpose of this part of the methodology is to provide an in depth analysis of the policy and regulatory framework that affects the access to crop genetic resources in agricultural markets. The focus of the analysis should be derived from the results of the VCA analysis indicating key policies and regulations that affect the market chain. Essentially the analysis will need to comprise a review of the existing national legal and policy regimes relating to access, exchange and sustainable use of plant genetic resources for food and agriculture (PGRFA in the case study country, followed by an assessment of how these policies impact the access to diversity in markets) building upon the findings of the VCA in this regard. This analysis will require a literature review and interviews with key stakeholders in the formal and informal seed and agricultural development and marketing sectors. The analysis should provide a broad overview of the legal, policy and institutional structures in the case study country that affect the content and quantity of seed supply focusing on the specific target crops of the case study.

Specific issues that should be covered in all of these analyses include the following:

- A review of the legal and regulatory regime in the formal seed sector, with a focus on the targeted crops should be provided. The analysis shall include, but shall not be limited to seed legislation (i.e. legislation on development, imports, evaluation, release, production and distribution of seed varieties).
- An assessment of any other national legislation or policy that may have an impact on the development of agricultural input and output markets. The analysis shall include, but shall not be limited to regulation/legislation on agricultural inputs, including legislation on agricultural land ownership, farm credit programs and input supply, as well as regulation or legislation affecting agricultural commodity production (e.g. price or quantity controls/support; deregulation under structural adjustment etc.).
- An analysis of the potential impact of existing legislation and policies on the content and size of seed flow in the formal and informal sectors and to the extent possible provide quantitative estimates of the supply of local and improved genetic resources through varying seed supply channels to seed insecure farmers.

- A review and analysis of pending policy or regulatory reform that may affect agricultural markets and potential impacts on access to crop genetic resources in agricultural markets. Identify areas of potential concern where policies, regulations and institutions are potentially impeding farmer access to seed and plant genetic resources in either the formal or informal sector.

D. Market and Household Sample Surveys

1. Measuring crop genetic diversity

a. Introduction

The project breaks new ground in seeking to estimate the amount of crop diversity and its distribution and structure in seed markets. The research design requires the development and application of methodologies to measure diversity in retail market outlets and relate it to farm levels of diversity. We are primarily concerned with within (intra) species (genetic) diversity measurement. However, the amount and distribution of genetic diversity of any crop in a production system will often be affected by changes in the number and types of crops grown and estimates of inter-crop diversity may also be important to understanding intra-crop diversity patterns.

The methodology is designed to obtain data that could be used to answer the following questions:

1) What is the diversity of selected crops and their varieties present in local markets and grown by farmers in surrounding areas (served by those markets)?
2) What is the stock and flow of genetic diversity provided by the crop and varieties traded in local markets and grown by farmers in surrounding areas?
3) What is the diversity of attributes provided by selected crops and their varieties traded in local markets and grown by farmers in surrounding areas? Attributes are defined as the characteristics of the crop and varieties that farmers select for, such as high yield performance, yield under risk, preferred consumption and use characteristics.
4) What is the genetic diversity present at molecular or biochemical levels in selected crops and their varieties and grown by farmers in surrounding areas? Does the extent

and distribution of this diversity have similar characteristics to variety and trait diversity?

In order to answer these questions, several concepts require definition.

Determining the physical availability of genetic resources involves describing the diversity of crops, of varieties within crops, and of genetic variability within and between varieties. Crop genetic diversity can be estimated from information on varieties, agro-morphological traits or from biochemical or molecular markers (or combinations of these). In all cases data is needed on the numbers (or variation in the case of quantitative agro-morphological characters) of different types and the proportions of these.

b. Establishing Varietal Identity

Establishing variety identity or, more accurately, the way in which variety identification works and the extent of an agreed common identification of varieties, is central to obtaining the required estimates of diversity. The number and amount of varieties can themselves be used to provide useful first estimates of genetic diversity. Provided that varieties are commonly and consistently identified, they provide information on the structure and organization of diversity in a production system providing the framework to link individual farm populations or seed-lots. Linking on-farm and market diversity data will also require information on variety identity and the extent to which market identification and on-farm identification are the same.

Within the project, we are concerned both with traditional varieties (landraces sensu Harlan) and modern varieties. The procedures adopted should ensure that information on both is obtained, and that whether varieties are properly classified as traditional or modern wherever possible. The problems of variety determination, especially for traditional varieties, are well known. Farmers may differ in the names they give to their varieties, using the same name for different varieties, different names for the same varieties or even no particular names at all. These problems become more common as one moves from community to community and are particularly acute for less common varieties. The procedures described in the Annex are designed to obtain a good picture of variety identity within a community. They can be expanded across communities in the research area. They involve focus group sessions with farmers, household surveys and discussions with key informants.

It is expected that variety identity may be treated more casually in markets than on-farm and create further problems for identification. It also needs to be recognized that, even when the same names are used within a marketplace and a community for agro-morphologically the same entity, there may be much more contamination or mixing within markets. In the first instance, similar procedures are proposed: focus group discussions with seed sellers, seller surveys and key informant discussions involving both expert farmers and sellers. These will need to be augmented with field trials to help resolve identity questions.

c. Principles for sampling

To answer the questions that the project seeks to address on genetic diversity, information on three different systems is required: on farm, in markets and the flow between markets and farm. This section provides a discussion of each of these, and their implications for sampling for information on crop genetic diversity.

i. On-farm.

The project will require estimation of the amount, distribution and structure of diversity within the chosen area(s) for the target crop(s). The approach for this draws on the experience of the on-farm project carried out by IPGRI. Data on numbers and identities of varieties maintained, areas of production and seed sources will be obtained in the household sample surveys. This will be complemented by focus group discussions and participatory rural appraisal to confirm variety identification (see below), traits of importance, uses, seed sources and crop areas in the community.

Seed samples will be needed for agro-morphological studies. Some estimates of variation in certain agro-morphological traits may be obtained from field visits during household interviews (if done at appropriate points in the plant growth stage) but this can only be done for traits of high heritability unlikely to show much genotype by environment G x E interaction and will need to be complemented by data from agro-morphological trials.

If at all possible these should be collected for each variety described by each household interviewed. Although this may seem to be excessive it does not seem possible to design an adequate stratified sampling strategy in advance of the results of the household interviews. The samples provide the necessary basis for confirmation of variety identity and studies on the extent and distribution of genetic diversity using agro-morphological traits or molecular markers.

The samples will need to be collected in such a way that they can be used to obtain appropriate measures of diversity. The IPGRI on-farm project suggested sampling 30 plants per household variety but this may be impractical for this project. We propose a bulk sample of each described household variety that is collected (this parallels the sampling procedure for market samples).

ii. Markets

Data on the different varieties sold and the amounts available will need to be obtained from vendor interviews during most of the seed selling period. An understanding of varietal identity – or, perhaps more strictly, a clear understanding of the units of management within the market for seed – will be essential. It will also probably be more problematic and difficult to obtain than for the on-farm situation but the suggestion is to follow the same general approach – a combination of individual seller interviews and group discussions.

Estimates of agro-morphological diversity and of agro-morphological characteristics may be especially important for understanding diversity in markets. Unlike the on-farm situation there is no possibility of investigators seeing much except specific seed, it may be the only way of checking identity. Again the data complement information obtained from sellers and purchasers on their perceptions of traits possess by materials in the market.

Seed samples will be needed for agro-morphological and molecular genetic studies. If at all possible these should be collected for each variety identified by each seller interviewed. As above, although this may seem to be excessive it does not seem possible to design and adequate stratified sampling strategy in advance of the results of the seller interviews.

iii. The relation between market level and farm level diversity.

The aim of the work is to investigate the relationship between genetic diversity in the market system and that on-farm. The design for sampling, the data collected and the experimental

protocols need to be appropriate to explore this relationship. This seems to involve three elements:

a) Determining flow from markets (e.g. as numbers of varieties and amounts sold as noted above) and use of markets by farmers from data collected in markets;
b) Determining the sources farmers use to obtain their planting materials from data collected through household surveys on seed sources for each variety and amounts obtained; and,
c) Including in the design a number of markets so as to allow comparisons to be made between them with respect to diversity present in markets and its distribution in surrounding areas.

An important element will be the relationship between varietal identity on-farm and in markets. This includes both the ways in which varietal identification is used on farm and in markets and the actual degree of agreement on names, identities and properties of the different units/varieties. The best approach might be to make use of expert farmer focus groups and data from the agro-morphological field trials.

The hierarchical structure of data collection (households →communities → marketshed) will allow for some consideration of differences in diversity at different scales. However, it is known that often there are substantial differences in the varietal diversity from year to year (e.g. in numbers of varieties, identities of common varieties and sources of materials). Future work to construct a panel of data on diversity in markets over time is thus highly desirable.

d. Data requirements

Three kinds of data can be obtained and used to estimate diversity at various levels and in various situations:

- Variety numbers and amounts – within and between households, communities, sellers and markets
- Agro-morphological variation for selected traits – within and between varieties, households, communities, sellers and markets
- Molecular (or biochemical) genetic data – within and between varieties, households, communities, sellers and markets

i. Varieties

The number of varieties and the quantities grown, available or sold provides the basis for estimating variety diversity in terms of richness and evenness and for comparing diversity present across households, communities, sellers and markets. Divergence between farms within a community, between communities within an area can also be calculated. Varietal diversity measures will allow studies of some of the factors (access, wealth, land holding etc.) that might be related to the distribution of diversity. Obtaining adequate estimates of variety diversity at each level should be a primary objective.

However, variety diversity will not fully reflect many aspects of the genetic diversity within an area. Varieties may be more or less similar, differing in a few or many characters and may differ significantly in the amount of genetic (agro-morphological and molecular) variation within or between them.

ii. Agro-morphological variation

Field trials (preferably carried out with the community) will provide information on variation within and between samples for agro-morphological traits of interest. These will include important highly heritable characters (e.g. seed colour, panicle type) and quantitative traits

(e.g. height, earliness). The information has three purposes. It complements farmer observations on important traits, it provides a way of confirming variety identities and allows the estimation of the amount of agro-morphological diversity on-farm and in markets and the way it is distributed between and within markets, communities, households and varieties.

Field trials should preferably be grown within the community using 3-4 replications over 2 years. It may be desirable also to have some trials at a local experimental station if this will allow more extensive data to be collected. The numbers of samples that should be collected has been noted above. However, not all will need to be grown in any one trial depending on the particular purpose of the trial. It is probably worth growing at least one trial in which market and farmer collected samples are compared.

iii. Genetic variation

The full answer to the questions posed by the project will probably require substantial genetic analysis, preferably using molecular markers. However, at least in the early stages, it has been suggested that resources for this will not be available. It is proposed that samples are maintained and used for specific questions once analysis of variety and agro-morphological data has been carried out.

Molecular markers will prove particularly useful for determining such questions as:

- The extent to which different samples of a variety (from sellers, households, locations etc) differ
- Effective population size
- The origins of specific populations or varieties
- The extent of differences between varieties
- The total genetic variation in a community

The numbers of samples are very large for genetic analysis. A first approach could be to adopt a sampling strategy that involved taking 5 seeds from each of an appropriate set of samples and estimating diversity using a suitably informative set of molecular markers (microsatellites, AFLPs etc). This could be used to determine e.g. whether markets had more or less genetic diversity than on-farm, whether samples of specific varieties contained the same genetic diversity on-farm as in markets etc.

e. Diversity indices

All diversity measures require data on the numbers of different types and the amounts of each type. These can be combined in many ways to provide alternative diversity indices which have particular properties and are useful in different situations. The two main factors taken into account when measuring diversity are richness and evenness. As species richness and evenness increase, so diversity increases (Hartl and Clark, 1997).

Richness is a measure of the number of different kinds of organisms, types or alleles present in a particular area. It gives as much weight to those types which have very few individuals as to those which have many individuals. Evenness compares the similarity of the population size of each of the species present. Two samples may have the same richness (number of species) and the same total number of individuals but different evenness. This is because the total number of individuals in one sample might be quite evenly distributed between the species. In the second sample, most of the individuals might be from one species, and only a few from the others. The second sample is therefore considered to be less diverse than the first (Hill, 1973; Hartl and Clark, 1997; Tothmeresz, 1995, Mouillot and Leprêtre, 1999).

For qualitative data (varieties names, qualitative morphological traits, molecular markers) two of the commonly used indexes are the Simpson's Diversity index and the Shannon-Weaver Diversity index (Table 2). The Simpson Index is particularly useful for on-farm variety data as it is relatively insensitive to misidentification of rare varieties. Generally, there is a consensus around identification of common varieties even where different names are used. However varieties which occur seldom and known to very few farmers may easily be misidentified or misclassified.

The partition of indices such as Simpson or Shannon Weaver provides a measure of the extent of divergence between farms within communities or between sellers within markets.

Table 2. Examples of Diversity Indexes

Index	Concept	Construction	Explanation
Count	Richness	$D = S$	S= Number of managed units of diversity
Margalef	Richness	$D = (S-1)/\ln A_i$	A_i = Total area planted or total population count or quantity over all units of diversity.
Shannon	Evenness, equitability, proportional abundance	$D = -\sum \alpha_i \ln \alpha_i$ $\alpha_i = n_i/N$ denoting N the sample size and n_i the number of observation in the category i. $D \geq 0$	α_i = population share occupied by the unit of diversity
Simpson	Proportional abundance	$D = 1-\sum \alpha_i^2$ $\alpha_i = n_i/N$ denoting N the sample size and n_i the number of observation in the category i. $0 \geq D \leq 1$	α_i = population share occupied by the unit of diversity

Source: Smale, 2006.

Where genetic data exists we can use Nei diversity statistics and Wright's F statistics to investigate population genetic structure and compare populations, varieties etc. (see Appendix B).

For quantitative data (quantitative agro-morphological traits) the coefficient of variation is commonly used together with appropriate analysis of variance procedures, $CV = s/\bar{x} * 100\%$ defining s as the standard deviation and $\bar{x}$ as the mean of the sample. The CV expresses the variability (or dispersion) of a sample relatively to the mean of the sample. A hierarchical ANOVA can be used to partition variation and identify significant differences in the different populations, varieties etc.

f. Linking measures of diversity to sustainable utilization

We are concerned with measuring crop genetic diversity in the context of sustainable utilization which necessitates collecting information and developing indices related to each of the three dimensions of sustainable utilization defined in section I above:

1) Private benefits to farmers via the consumption and production "services" that genetic resources (in interaction with human and environmental factors) provide to farmers
2) Maintaining genetic option values (reducing genetic erosion)
3) Increasing resilience (reducing genetic vulnerability)

A brief discussion on how the genetic diversity measures may be linked to the notion of sustainable utilization follows below.

i. Measures of crop genetic diversity associated with private benefits to farmers

One key source of information on this aspect of sustainable utilization will be derived from information on variety attributes from both improved and landrace varieties collected at the farm and vendor level, together with information from an agro-morphological survey conducted on seed samples collected at the farm and vendor level. The agro-morphological information will be used to support the descriptive information farmers provide about variety traits, such as resistance to drought and disease, early or late maturing etc.

The set of attributes farmers demand from varieties (based on what they are currently planting as well as contingent questions on varieties they would like to plant) will be compared with the set of attributes associated with the varieties supplied in markets. The market set of attributes will be derived from focus groups with vendors. Assessing information about the attributes demanded by farmers and what they receive from markets is also important. We expect that there will be some divergences between both the attributes and the information about the attributes that farmers demand and are provided in the market.

It is important to note that this measure of diversity will not be linked to traditional or improved varieties per se, but to the entire set of attributes provided by both traditional and improved varieties. Thus we may find that utilization patterns with high number of improved varieties that provide a wide range of attributes are considered more diverse than utilization patterns based on a small number of landrace varieties.

ii. Measures of crop genetic diversity associated with the reduction of genetic erosion.

Genetic erosion is of primary concern for landrace varieties. We are concerned to maintain viable populations of landrace varieties in the field, particularly in areas of primary or secondary diversity for the crop. Our site selection was designed to allow us to gain insight into issues of genetic erosion since we are working on target crops in areas of primary or secondary diversity. However, we do need to be clear that we are not able to get a direct estimate of this benefit as we are not planning on doing time series studies – at least not at this point. This benefit is measured by a change over time and our data on diversity will be cross-sectional. The normal way around this problem has been to use some kind of spatial surrogate in which areas with low diversity (in terms of area in landrace varieties) are compared with those of high diversity and assume this reflects a temporal possibility which is the way we are approaching the issue in this study. Analysis of relationships between richness, evenness and

divergence can be used to explore the extent to which the crop and variety in the area/market of choice are rare.

iii. Measures of crop genetic diversity associated with reduced genetic vulnerability

Measuring this benefit requires spatial measures of the degree of variety richness, evenness, and divergence (the probability than any two samples from a particular source will be identical) associated with the utilization pattern. These are simple, but powerful, and easily understood. The benefits from reduced genetic vulnerability are realized at varying scales so we will need to develop indicators at the farm and community level at a minimum.

2. Quantitative Market Surveys

There are three instruments for conducting quantitative surveys at the retail market level:

- The market observation survey
- The key informant survey
- The vendor survey

The key informant surveys at the retail market level can essentially be thought of as parts of the seed value chain analysis. It may be the case that the key informant is a member of the research team – as in Mali. However, the questions are more standardized for the VCA – e.g. we need to have a standardized set of information to proceed with the quantitative surveys of the markets and vendors. In these surveys we will be collecting information on various aspects of the market structure, conduct and performance, including measuring the diversity available via questions and seed sampling.

a. Sampling strategy: selecting retail market outlet sites for sampling

Study sites were selected to be located in areas with a crop that was important both for its local diversity and to food security, and where markets are an important source of seed for the target crop. In this section we discuss how to select the retail market outlet sites to focus on within the selected study site.

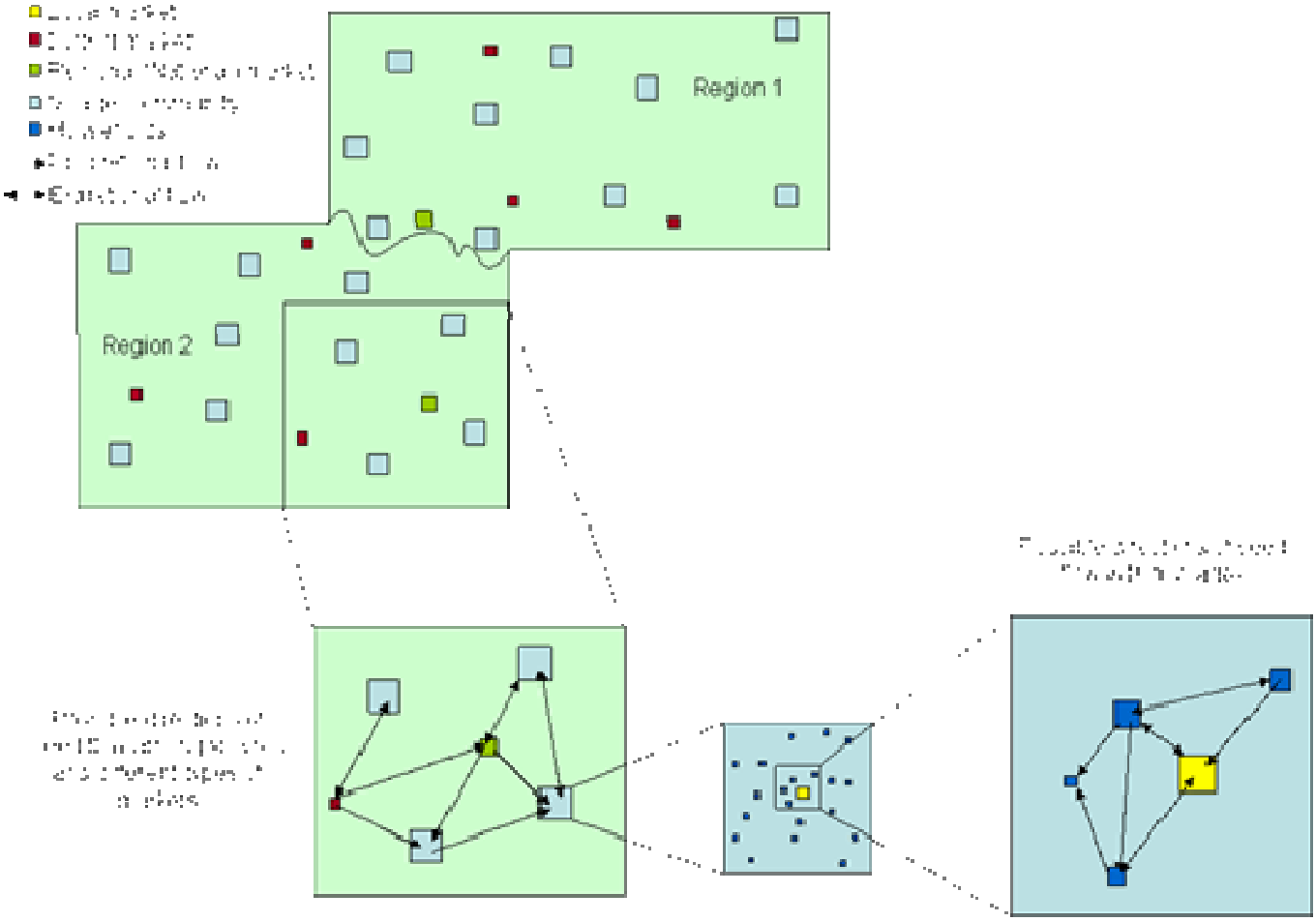
Two key principles guide the selection of market sites for sampling:

> i) Market sites should be purposively selected to have significant differences in crop genetic diversity (e.g. markets within and outside areas with seed system interventions likely to impact market functions) Market site selection should be done using information from the value chain analysis on the characteristics of the "marketshed".
>
> ii) All other factors that could have an impact on farm level diversity and welfare should be kept as constant as possible – particularly important to try to select markets in similar agro-ecological zones, and also population density should be similar

1. Defining a "marketshed"

We define a marketshed as a geographical area and associated population that has real or potential trade relationships with a market center. Within the marketshed there may be several market outlets of varying size and scope, although usually one is dominant in terms of size and function. We use the term marketshed to give the sense of a system or network of market flows within a given area (e.g. it is not just a zone with some markets located in it, but rather it is a trading network with links between market outlets). In the course of the value chain study, the structure of the marketshed should be revealed –e.g. the links and nodes between the market outlets within a marketshed as well as the size and nature (variety characteristics, sources) of seed flows in and out of these markets. The diagram presented by IPGRI at the 2006 workshop and shown in Figure 5 below gives one example of what a marketshed for seeds could look like:

Figure 5. **Hypothetical distribution of markets and villages and seed flows in a marketshed**



2. Control and Treatment Marketsheds

In our study, several of the case studies are using the presence of a project or intervention (e.g. emergency seed relief, diversity fairs) to define the marketshed based on an assumption that the intervention has an impact on the access to crop genetic resources in markets located within the marketshed.

Think of the seed market/system intervention as the "treatment." The effects of the treatment are felt to a greater or lesser degree by farmers within a set of villages. At least some farmers in these villages will participate in seed and product markets. These villages are located in a market zone and the zone will likely have multiple markets. There is an imaginary, imperfect boundary around the "treatment zone." In Mali, Kenya, and India, we have rough idea of where this boundary is (Figure 16).

Now think of the same structure in an area in which the impact of the intervention cannot be felt. This zone is the "control zone." In the case of seed, as compared to a poverty impact program, it will be hard to separate control and treatment groups. This is because seed moves among farmers, sometimes at great distances (Edmeades, Jackson and Smale, 2007). Often the

control and treatment groups must be sampled purposively, to ensure that they "match" regarding fundamental parameters expected to influence impact variables (Ravallion, 2005)

From a sampling standpoint, to counteract intra-class correlation, it is better to sample more villages and sample fewer household per village, though it raises logistical costs. From a data entry and analysis standpoint, a self-weighting sample (constant sampling fraction, which implies variable sample sizes within villages) is more convenient. Obviously random samples should be drawn.

To avoid the problem of "program placement" bias, it is important to know, understand and document: 1) the reasons why the intervention was placed where it was placed; and 2) to hold constant (between control and treatment zones) observable factors that are expected to play a major role in the impact of the intervention, such as presence of another development program, ethnicity, agroecology, and overall level of development of market infrastructure. There is an imaginary boundary around both treatment and control zones (Figure 6).

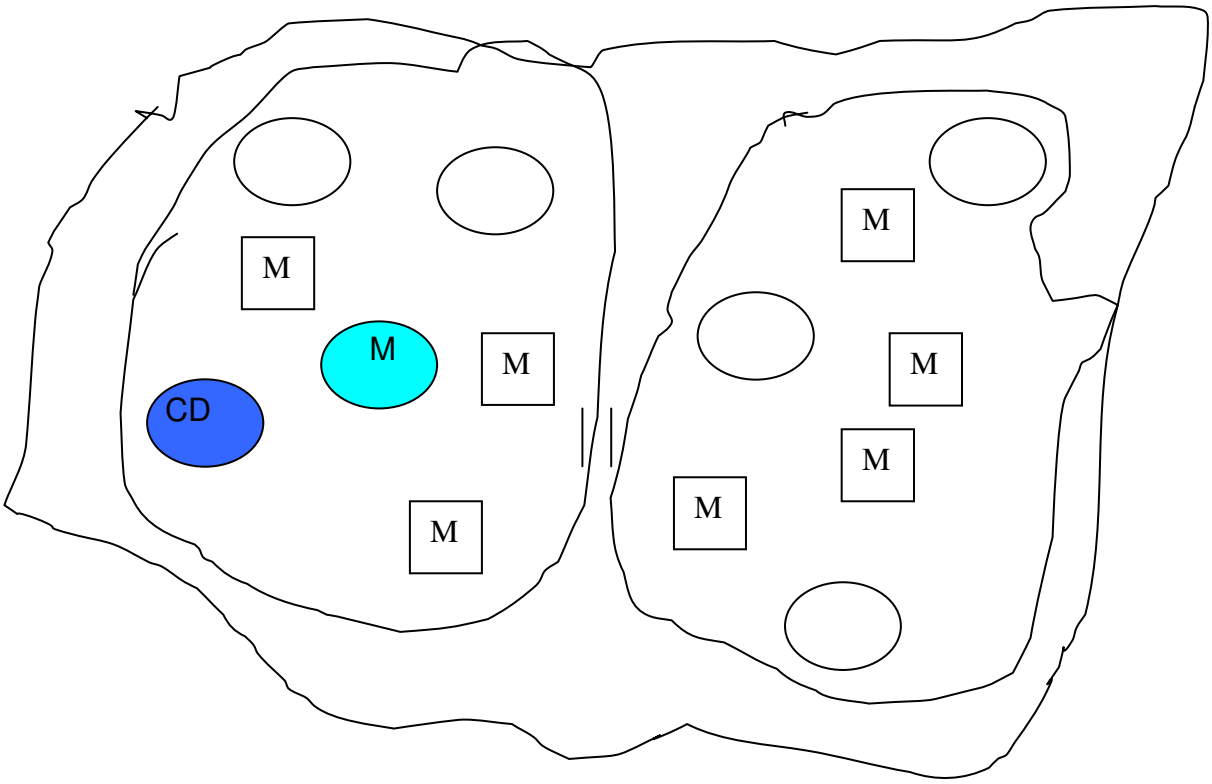
The Mali project provides one example. In the Mali project, the markets are weekly markets, and they are all node 1 (retail) markets, although they do have different characteristics that will be observed and documented (using the market observation and vendor surveys), as well as the value chain analysis. The intervention that defines control and treatment zones in the Mali case are: 1) diversity field fora, and 2) diversity fairs. See Figure 6 below. Within both control and treatment groups, we have farmers who participate in markets to purchase grain as seed, and those who do not. This implies that we can test a series of hypotheses which are laid out in section 4 below.

The India project provides a different example. In the Kolli Hills, the M.S. Swaminathan Foundation has intervened in a number of ways to develop the market channel for minor millet products. In the Kolli Hills, both control and treatment groups have been sampled. Still, farmers are autarkic with respect to seed. In the Dharmapuri District, the major millet-producing district of Tamil Nadu in the nearby plains, there are well-developed markets for grains but not for seed. There are no specific interventions to support marketing of minor millets in the plains. In a qualitative way, the Kolli Hill situation, as compared to the plains, provides a comparison of finger millet diversity in the absence and presence of markets. Seven markets have been sampled in the Dharmapuri District, so that the effects of participation in markets on farm-level diversity, as well as the effects of variation in market diversity on individual farmers, can be tested.

In Kenya, two types of interventions have been undertaken for pigeon-pea. One is a non-market intervention of direct seed distribution. The other is a market-based intervention, consisting of producer market groups and community seed production for sale. Again, seven markets have been sampled. In this case, it may be possible to either test the effect of the market intervention on market diversity compared to no intervention or compare the effects of market and non-market interventions.

Figure 6. Diagram of control and treatment marketsheds

Ovals are villages and boxes are node 1 weekly markets. The two interventions are shaded in blue (diversity field and seed fair). Note – the overall zone should be as homogenous as possible in terms of agro-ecology and population density.



3. How many markets within the marketshed should be selected for the sample?

This has been a major source of discussion – essentially the geneticists want more, the economists think less is sufficient. With limited budgets a critical issue is deciding how much effort to put at the market vs. household level. The sampling strategy adopted should allow for comparison between on-farm diversity and market diversity and for understanding the relationships between them. Variation in the amount and distribution of diversity may be associated with differences at household, village or market levels and data needs to be collected in ways that these possibilities can be tested. Since market characteristics are of major importance, it is suggested that enough markets (at least 5 across all marketsheds) are included to allow comparisons between them to be related meaningfully to differences in diversity together with a number of different communities/villages. As noted in the preceding section, Kenya and India include a sampling of 7 markets. Initial results from the case study sites seem to indicate that including at least 5 and up to about 10 markets is reasonable to cover given the amount of work required for conducting the market surveys and the numbers of households we are expecting to survey (around 400 per case study country). This is an issue that will need revisiting as initial results become available.

In considering the minimum number of desired observations, it is important to think about the units of analysis. For the markets we are likely to have observations on the level of diversity (and perhaps other measures of access) at more than one period of time, and most of the

quantitative measures of market access to crop genetic resources (and other measures of market conduct and performance) will be generated from the vendor survey. For some variables we can have an N based on number of vendor observations over time- which of course will be much higher than the number of market observations – or the N on variables created as aggregations of vendor observations to a market level (e.g. total number of varieties sold in a market over a season). The same could apply at the household level – we will be collecting information at a plot level as well as a household level. For some variables the unit of analysis will be the plot which will give higher Ns than the household level data.

4. Timing

The vendor survey will take place over the period at which there is seed in the markets. Information on the degree to which there is periodicity in seed sales (in what is sold and prices) should be collected in the key informant analysis. Where there is evidence of clear periodicity in seed sales, we suggest weekly visits, planned to cover at least 70-80% of the seed selling period. These are most likely to yield good data on the types of varieties sold, amounts and prices over the market season. If there is no evidence of periodicity then information about prices, volumes and variety types could be collected with a one time interview, however it could still include questions on maximum, minimum and median values over time.

b. Sampling Strategy: Sampling Vendors for Information and Seed-lots

The seed vendor survey is the primary instrument used to collect quantitative data on the genetic diversity, information and prices that are present in the market. We will be collecting both information and seed samples from the seed vendors and we will be gathering data from them (primarily on prices and quantities sold) repeatedly over the relevant sales season.

In some of the case study sites there are few vendors within a market, or the vendor and the market are the same, and therefore the vendor survey will end up being a census. However when there are several vendors in a market (say more than 10) then it will be necessary to draw a sample. Ideally, the sample of vendors should be drawn to provide a sample representative of the key components of market access to crop genetic resources: e.g. the diversity available, the information about of the diversity available and the prices. In collaboration with other groups involved in the research, a decision will be needed on whether to visit the same sellers each week or to select randomly vendors during each visit.

Box 3. gives a description of the different types of vendors present in markets for the Kenya study.

Box 3:
Types of vendors in the market in Makueni, Eastern Kenya,

Five types of vendors were identified that fall into two main groups: those with permanent structures and those that trade in open-air:

A. Vendors with permanent structures

These vendors have their own permanent structures (building-rented or owned and other structures needed to conduct the trade) in the market premises and operate throughout the year. These include:

1. Agro-vets (dealers of improved crop varieties (seeds), pesticides and fertilizers)
2. Grain stores (exclusive stores that store grains of different crops who sell seeds during times of heavy demand especially before planting)
3. Local shops or shop keepers (Shops that store grains also apart from other items of grocery ranging from bread to batteries)

B. Open-air Vendors

Vendors in these categories do not have any permanent structures inside the market premises; they pay market fee to enter the market for trading on the day of the market and they trade from places allocated (by the municipality officials) for them to do so. These include:

4. Market Hoppers (traders whose major profession is selling grains or seeds depending on demand and they hop between various markets in the community)
5. Farmer-Traders (these are farmers who bring their own surplus grain or seeds to the market and also occasionally purchase from their neighbors and friends in the community for trading the open-air weekly markets)

C. Characteristics of farmer traders and market hoppers

These are mostly farmers and bring their own produce mostly to trade in order to buy other consumption goods of interest such as vegetables, oils, clothes with the money earned. Some of them do sell their produce directly to the grain stores or local shops in the markets if time is constraint for them. Other women (90 percent of these traders are women) conduct business throughout the day or till they sell whatever they brought.
In the semi-arid regions of Eastern Kenya, just before planting season –which is Oct 15th – November 1) or immediately after a single shower, the following markets days would be very active with these kind of sellers.

These women traders can be characterized as: i) *one-time traders* i.e., they come to market with their surplus only for a single market day before planting or immediately after the shower just to make use of the immediate seed demand among the farmers and to make quick profits(seed prices are double than grain prices) or ii) farmers who bring seeds *continuously for all the 3-4 weeks* during the planting season every year (both own plus buy from their neighbors and friends) and sell it in the market, iii) farmers who also bring *pigeon pea along with vegetables* and sell it in the market over a period of time, iv) traders whose main activity is to trade from one market to another (within 5-6 markets in the community), hops from one market day to another *continuously*.

There is more than one way to go about sampling vendors. We have two different examples from the Kenya and Mali cases, as well as a third possible suggestion. The best method to adopt depends on the context. In Kenya a two-stage procedure was used meaning that the sample design for vendor lots is a form of stratified, two-stage cluster design. "Strata" are vendor types. The first stage sampling unit for the vendor lot is the vendor. The second stage is a census of all types of seed sold by the vendor. The two stages are described in Boxes 4 and 5.

Box 4:
Vendor sampling protocol for Makueni, Eastern Kenya,

Agro-vets, grain stores and local shop vendors

1. A detailed inventory of the number of agro-vets, grainstores and local shops present in each of the market locations (to be sampled[2]) in the Makueni district has been prepared. From the list, it has been selected the vendor shops selling pigeon pea along with other grains or crops[3].
2. The total number of vendor shops dealing with pigeon pea (either as grain or seeds) in the 8 market locations is 160 in number. These shops are ranging from small, medium to large in terms of their transactions and also number of crops they deal with.
3. For the sampling purposes, all of the 160 vendor shops (belonging to all the three types of vendors including agro-vets) are included.

Market hoppers/farmer traders

1. On the "market day" all the open-air traders who brought pigeon pea solely or with other crops were counted.
2. To get the maximum number of participants, the traders were counted between 11.30-12pm (peak time of trading).
3. All the traders were interviewed in market locations where the number of traders was less than 15.
4. In the markets where more than 15 traders were present, say 30 (pigeon pea—our focus crop), we counted them T1 to T30. Usually all these traders are seated next to each other in an open-air market at a certain part of the market that sells food grains/seeds.
5. Out of these 30 daytime traders, 60 % (say 18 of them) were interviewed.
6. These 18 traders were picked randomly using the following procedure (this procedure was described as the investigators at times couldn't understand the definition of random numbers or generating them). It was told to the investigators to choose ,from the first 10 traders (from T1 to T10), the "ODD(even)" number traders; for the next 10 traders (T11-T20) they were told to choose the EVEN(Odd).
7. This procedure was conducted for two market days in the same market.

[2] The market locations selected for the purpose of the study are based on seed intervention and non-intervention areas. The names of the market locations are: Wote, Kathonzweni, Kalawa, Kasikeu, Sultan Hamud, Emali, Mulala.

[3] Except for Agro-vets though most of the agro-vets do not sell pigeon pea seeds(less than 1 percent) we would like to include them for our survey purposes as they are the only formal sources of seed supply (for other crops such as maize and beans) which could form potential sources of seed supply in future.

As noted in Section B on Measuring Diversity, seed samples will be needed for agro-morphological and molecular genetic studies. If at all possible, these should be collected for each variety identified by each interviewed seller. Although this may seem to be excessive it does not seem possible to design an adequate stratified sampling strategy in advance of the results of the seller interviews.

The samples should be collected in such a way that they can be used to obtain appropriate measures of diversity. Before seed lots can be sampled, the phenotypic taxonomy used by farmers to characterize the unit of diversity they manage (variety ideotype) must be known through focus groups and previous scientific study. The unit of diversity could be a "named variety" but is not necessary. Often it is a name and a set of distinguishing characteristics, and set of agronomic and consumption attributes. There, is of course, a relationship between these and genotype. Direct purchase of reasonable quantities of seed on two separate occasions would seem appropriate (as might be done by farmers seeking seed for an average sized plot).

The ability of traders to distinguish varieties affects the sampling strategy. Previous research in India indicates that traders were able to differentiate types of millet crops, according to characteristics of the grain, whether it was a modern variety or farmer variety, but not by variety name (Nagarajan and Smale, 2005). However, this is not the case for pigeonpea in Eastern Kenya. No more than three types per vendor were found across the five different vendor types (see Box 3). While traders reported some differences by color and grain size, visual examination of the grain (peas) in the sacks showed no appreciable differences. Among those who sell in the open air or the grain stores, the names were reported for farmer varieties only when sold by a farmer-trader or by a vendor who learned it from the farmer from whom it was obtained. This was important in determining the sampling strategy for the seedlot (see Box 5)

Box 5: Seedlot sampling protocol for Makueni, Eastern Kenya

1. In the vendor questionnaire questions about varietal identity of pigeon pea sold by different types of vendors in the market (as described by the vendors and not imposed by an outside investigator's perception) were included.
2. These questions are followed by the detailed description of 'product characteristics' (section 5 in the vendor's questionnaire) describing the pigeon pea lot the vendors sell or store in their shop.
3. 250 grams of pigeon pea were sampled for each type of pigeon pea and from each vendor. If, for example, the farmer trader said "I brought three bags of pigeon pea to the market, one is Katoli(farmer variety) and other is Kionza (improved) and the third bag I sell I do not know the name of, but I think it is a mixture of unknown varieties," then it was asked to the trader to describe the characteristics of all the 3 seed lots that she has brought to the market. In this case, the investigator is asked to buy 250 gms of each of the three lots.
4. The investigator is then asked to write the following details in the brown bag: investigator name, location of the market, type of the vendor, date of collection, the name or variety (as told by the trader), any other significant details regarding the lot.

In sampling households and measuring their characteristics, a common concern is intra-cluster correlation. That is one reason why it is recommended to partition a sample more widely across villages with fewer households per village. Here, however, there is not evidence of a source of intracluster correlation because the characteristics of the vendor seed lots are not

correlated with the vendor as much as with the environment of the grain or seed. This sampling approach differs from the one for Mali described in Box 5bis.

Box 5bis: Sampling seed lots in markets: the Mali protocol

It is important to have someone on the local team who is knowledgeable and who will take responsibility for the seed samples, their storage, and their growouts. However, make sure that the social scientists and crop scientists are together when the sampling is done! Otherwise, the link between the genetics and the household farm might be lost.

Here are the steps followed in Mali:

1. Go through the market recording the number of vendors of millet/sorghum grain/seed in the market on that day.
2. If there are 25 vendors or less, the sample will include all of them.
3. If there are more than 25, take only 25 selected at random, regardless of the number of vendors.
4. Take a paper with a blank matrix that shows grain/seed type (or variety) in columns by vendor (v=j, where =1,…..k) in rows (See Table 3 below). Note: type is defined by the characteristics that vendors use to distinguish grain (seed) they sell. It may be that type=variety, so that variety identity is preserved. If not, we will need to elicit vendor discriminating characteristics.
5. Go to each vendor in the market on that day and list each new variety across the top of the column, ticking downward each time a vendor sells a variety. You will have a census of all varieties sold on that day by all vendors.
6. Calculate ln (vjm+1)=vjm. The number of grain (seed) samples to take per market per type is equal to the natural logarithm of the number of plots sown + unity.
7. Purchase 2 bowls of millet per vjm and 2 bowls of sorghum at the going price.
8. Administer the vendor questionnaire. Buy them a cold drink.
9. Do this from grain (seed) available on the market before the planting season in 2007, or two visits per market. For example, for 4 markets per market zone (total of 16 markets) and 25 vendors per market with an average of 3 varieties per vendor and 2 market day visits.

Table 3. Census of all varieties sold by all vendors in the market

Vendors/varieties	**Variety A (yes/no)**	**Variety B (yes/no)**	**Variety C (yes/no)**	**Total**
Vendor 1				Total number of varieties sold by vendor 1
Vendor 2				
Vendor 3				
Total Number of vendors selling each variety	Total number vendors selling variety A	Total number of vendors Selling B		Total (total number of varieties sold by all vendors)

The Mali approach is adjusted for dominant and rare types by "smoothing" the number of lots with the logarithmic function. In doing so, it reduces the number of lots that need to be characterized and it enables good analyses of the between- and within-type diversity. In the case of Kenya, there are several types that are mixed or unknown so that the diversity analysis may be more challenging. On the other hand, the diversity analysis may be helpful in identifying where mixtures and unknown types relate to known types. These statements, however, need to be tested against expert (geneticist) opinion.

A third option may be to stratify the vendor sample based on information gathered in the key informant survey. This method is attractive for cases where there are significant variations among vendors in terms of products and pricing for the target crop varieties. It requires having prior information on the number of vendors in a market and the degree to which they vary over the relevant seed selling season, the distribution of varieties or seed lots among them (e.g. are they all more or less selling the same thing, packaged in the same way?) and the degree to which there is price variation amongst the vendors for a comparable product. These are all issues which should be addressed in the key informant survey. The key informant survey should give a quick census of what is being sold in the market, for example how many bags being sold and how many different varieties. This census should be structured to provide the necessary information for sampling vendors – but can also be an important part of the study output itself in producing an assessment of the distribution of market share amongst vendors and how it compares with the number of varieties being sold.

A two step process is needed to do this. The first step is to fill out a table such as Table 3. This table gives a census of all the varieties that are being sold by all vendors in the marketplace. It does not however, give any indication of the market share of the various vendors or the relative volumes of varieties being sold. This is addressed in Table 4.

In Table 4, information from Table 3 is used, together with information from the key informant survey on the relative market share of the various vendors, in order to group the vendors by market share and number of varieties handled. Vendors are grouped by whether they are high volume (in terms of seeds sold of target crop over the entire season) and whether they sell many or few varieties. Data for the market share comes from the key informant survey, while data on number of varieties handled by vendor and by vendor group, comes from the data collected to make Table 3. In general, it would be expectedthat repeated vendors (those that sell over the entire season and who have a fixed market facility) are most likely to fall into the high volume category, while individual traders that only sell sporadically are likely to fall into the low volume category.

Table 4: Categorization of Vendors by Market Share and Number of Varieties Sold

Volume/Number of Varieties sold	**Many varieties (>3)**	**Few varieties <+3)**	**Total**
High volume (~> 20% market share)	**A** (Number of high volume vendors selling many varieties) An example of a high	**B** (Number of high volume vendors selling few varieties) Likely examples here are representatives for MVs (for	Total number of high volume sellers

	volume seller who has many varieties could be market hoppers in the Kenya case.	example hybrid corn dealers). Could also be seed shops in Kenya case	
Low volume	**C** (Number of low volume vendors who sell many varieties) (Total number of varieties handled by high volume sellers) This category is likely to be sellers of seeds for niche markets – perhaps farmer-sellers. Example may be from Kolli Hills in India	**D** (Number of low volume vendors who sell few varieties) (Total number of varieties sold by all vendors in the category.) This is the category where we'd expect the largest numbers of vendors to fall into. Here it is also important to get a sense of how much variation there is in the varieties sold by each – so as a group do they represent a wide range of varieties or are they all selling pretty much the same thing. This is likely to be petty traders participating in open-air markets.	Total number of low volume sellers
Total	**Total number of vendors selling many varieties (give an estimate of total number of varieties included in category)**	**Total number of vendors selling few varieties (give an estimate of total number of varieties included in category)**	

From Tables 3 and 4, we can calculate 3 statistics for each of the 4 vendor groups defined in Table 4: the market share, the vendor share and the variety share that each box represents.

For Box A, we expect market share to be high, vendor share low and variety share high.

For Box B, we expect market share to be high, vendor share low and variety share low.

For Box C, we expect market share to be low, vendor share low and variety share high.

For Box D, we expect market share to be low, vendor share high and variety share is uncertain (while it may be low for any individual in the group it may be high for the group as a whole).

Ideally, the design of the vendor survey should be purposive using information from Tables 3 and 4. Vendors from all groups should be sampled based on market share and variety share. This is a little complicated, but ultimately drawing the sample should depend on information from the key information survey on how much variation we would expect to find among each of the 4 groups in their pricing behavior, variety labeling and information transmission- e.g. the two other dimensions of access to crop genetic resources in markets aside from crop genetic diversity.

c. Data to be collected in the market and vendor surveys

Farmer access to CGR in markets is determined by the availability, affordability, timing and information about seeds and their genetic content in local markets. The quantitative surveys are designed to obtain information on these dimensions. One of the primary innovations in this study is a focus on genetic diversity in the seed supply system, and thus as a determinant in the farm-level access to crop genetic resources.

An explanation of how we are proposing to define each aspect of accessibility to CGR in local markets is as follows:

- **Availability**: Defined as the quantity and diversity of the seeds of adapted crops and varieties on supply in a market, including the range of genetic resources they embody and the attributes they provide. (See section on measuring crop genetic diversity for more details.)
- **Timing**: This is essentially part of availability – that seed and the genetic resources of seed are available in the market at the time they are needed by farmers.
- **Affordability**: A key issue here is that the benefits or returns associated with the seed and its genetic content are higher than the costs of obtaining the material through the market. The prices of seeds and their genetic services, and other complementary input prices, will affect this cost.
- **Information and other transaction costs**: This affects all other categories. The degree to which a farmer will benefit from available and affordable seed in local markets depends on whether they know that the seeds are available, if they have information about the genetic content and/or quality of the seed, and the potential adaptability and interaction with the farmer's micro-agroecology. The difficulty and cost of obtaining this information is an important determinant of farm level access to CGR.

In addition, survey data and secondary observations on general market characteristics such as entry and exit costs, numbers of buyers and sellers, regulation, and infrastructure, and detailed data on seed or other planting material availability, prices, and information for the crop(s) of focus and its substitutes will be collected. Also, it will be collected information on the characteristics of the exchange (where, with whom, price, quantity, timing, etc.), of the market (number of other buyers and sellers, products, location, variety, frequency, etc.) and the physical and informational transaction cost constraints.

The data from the market surveys should allow us to develop the following set of descriptive statistics on seeds:

i. Number of varieties (local and improved) available
ii. Distribution of varieties available – to what extent are rare varieties handled?
iii. Types of attributes available from varieties offered in market
iv. Total quantity of seed flowing thru outlet/season
v. Distinction/mixing of varieties by outlet
vi. Information on genetic content by outlet
vii. Mean and variance of price by variety by outlet

Market structure and access to crop genetic diversity

i. Vendor characteristics (eg. ethnicity, gender, size, type of outlet) by:
   Number of varieties and packaging by variety (mixing or separated)
   Attributes of varieties provided

Type of varieties (local, improved)
Source of seeds (numbers, type)
Pricing
Contract/client type served

ii. Relationship between vendor dominance in markets and genetic diversity (e.g. Table 4).

d. Administering the quantitative market survey instruments

Templates for the market observation, key informant and vendor surveys are given in Appendix III.

3. Household sample surveys

a. Household sampling strategy

The approach we will adapt in several of the case studies has been developed and implemented by economists at the World Bank in order to evaluate the impacts of programs on poverty. The following paragraphs summarize some of this literature and basic terminology.

The household sampling strategy is linked to the market sampling in that the households should be located within the marketshed. Essentially we want to define a group of households that live within sufficient proximity to one or more markets to allow participation to be feasible.

The fundamental problem with evaluating the household level impact of market access to crop genetic diversity is defining the "counterfactual." The counterfactual is the situation of participants if they had been located in a marketshed with identical characteristics, but with differing levels of access to crop genetic resources in markets. This is usually unobservable.

To overcome this problem, social science researchers attempt to recreate an experimental situation as closely as possible to simulate the counterfactual. We employ a "control group" that is composed in such a way that it represents as much as possible the "treatment group" with the exception that its members did not participate. Evaluation is accomplished through the application of various tools, each of which has advantages and disadvantages, as shown in the table below.

Table 5. Methods for evaluating program impact

Method	Disadvantages
Randomized	• political and social feasibility • criteria of donors • selective non-participation • sample attrition
Matching	• non-matching of unobservable characteristics, leading to statistical bias • propensity scores depend on the model chosen by the researcher for representing the decision to participate
Reflexive	• ignores changes induced outside the program
Double-difference	• sample attribution
Instrumental variables	• difficult to identify variables that influence participation but not the indicator, considering participation

Source: Ravallion (2005)

With the randomized approach, errors are related exclusively to sampling, and can be reduced by drawing a larger sample. There is no statistical bias. With matching, the composition of groups is checked by measuring their characteristics and estimating propensity scores, calculated to justify the inclusion of each individual in each group. Reflexive methodologies evaluate the same group before and after the program, but ignore changes induced by factors outside the program—leading to attribution problems. The double-difference approach compares both treatment and control groups, before and after. The instrumental variable approach relies on econometric methods to separate the effects on poverty of participating in the program from other factors that influence poverty. The instrumental variables are those that influence participation but not program impacts, considering participation. Identifying such variables poses a challenge in applied research.

The sources of bias in impact evaluation are derived from the location of the project or program ("program effect"), identification of program participants ("selection effect"), and differences in observable and unobservable characteristics between the two groups. Differences between the two groups with respect to observable characteristics occur when, in statistical terminology, the two groups have no "common support" which can be reduced by adjusting the samples through matching and propensity scores. The bias in unobservable characteristics is often addressed through employing the double-difference technique.

In this project, there is an opportunity to improve upon each individual approach by combining several approaches. The matching, double-difference, and instrumental approaches

can be used[4]. Matching is used to enhance the suitability of the sample. Box 6 describes the household sampling scheme being used in Mali

Box 6.
Household Sampling Scheme in Mali

A self-weighting sample of 450 farmers has been drawn, with approximately 150 per site (3), and 75 in each group (control/treatment). At least 3 villages will be sampled in each of the groups, each set of three serving a different set of markets. In the treatment group, not only participants but non-participants will be sampled, since we know that there will be spillover effects on non-participating villagers and on villages where no "champs de diversité" (CDs) have been established.

Selection criteria for villages are: 1) location in a) San, b) Douentza, c) Tassiga sites; 2) villages within the areas affected by the FIDA development project; 3) presence of same NGOs; 4) same ethnic representation (e.g, in Douentza, one village Sonrhai, Dogon, Bambara); 5) villages within each group use the same weekly markets; 6) villages in the treatment group do not use the same weekly markets as villages in control group.

Another criterion to use in calculating propensity scores is that the villages have similar average levels of: 1) area cultivated per production unit; 2) farm machinery and equipment; 3) size of livestock herds; 4) number of employees; 5) education.

b. Sampling seed-lots from households

Before seed lots can be sampled, the phenotypic taxonomy used by farmers to characterize the unit of diversity they manage (variety ideotype) must be known through focus groups and previous scientific study. The unit of diversity could be a "named variety" but this is not necessarily the case. Often it is a name and a set of distinguishing characteristics, and set of agronomic and consumption attributes. There is of course a relationship between these and genotype.

1. Use sample design of household survey
2. Stack all the plot files, so that all of the plots of all varieties of millet and sorghum planted are listed.
3. Run the frequencies by variety (count all plots planted to each variety per village, or nij where i=variety and j=village).
4. Calculate ln (nij+1)=sij. The number of seed samples to take per village per variety is equal to the natural logarithm of the number of plots sown per village per variety + unity.
5. Explain the sampling procedure to the farmers and what the samples will be used for. There will be growouts in another site, but we will share all of the results of the analysis with them. The farmers will be paid the cost of the seed at the current grain price. The farmers must keep ALL the seed they select from the sample plot apart from other seed or grain.

[4] The reflexive methods might also be useful for stakeholders in their self-assessment, but without the application of economics methods.

6. The team will then select the 30 panicles/plot, selected at random.

c. Data collected on household surveys

We would like to explore the relationships between market access to crop genetic resources and the household decision to participate in seed markets, and the impact of market participation on farm level utilization of target crops genetic resources as well as measures of well-being and on farm crop diversity.

In order to do so we need targeted information from households on a number of issues

i) Measures of household welfare
ii) Utilization of genetic resources
iii) Crop attributes selected
iv) Market participation
v) Socio-economic characteristics of the HH

i) Welfare Indicators

There are many different ways to think about household welfare, and some are more likely than others to be affected by utilization and on farm diversity. Several studies have indicated that on farm diversity is negatively correlated with one measure of welfare – e.g. income, but now we'd like to explore whether the story changes if you look at a broader range of welfare measures. The two we are particularly interested in are dietary diversity and resilience to shocks. A brief discussion on each follows below.

Dietary diversity: Dietary diversity defined as the variety of foods eaten within and between food groups over a reference period is thought to be associated with a higher quality diet and thus better nutrition (Ruel, 2003). Research in developing countries has confirmed a consistent positive association between dietary diversity and nutrient adequacy (Ruel, 2003).

Dietary diversity can thus be considered an important measure of welfare. There are two levels of hypotheses to develop around dietary diversity:

A) Farm households who have a more diverse set of varieties of the target crop are likely to have lower levels of dietary diversity AMONG food groups. This hypothesis goes back to the notion that on-farm infra-specific diversity is negatively related to wealth. Assuming that dietary diversity is positively associated with wealth and infra-specific diversity is negatively associated with wealth, we would expect a negative relationship here.
B) However, farm households who grow a more diverse set of varieties for the target crop are likely to have higher levels of dietary diversity WITHIN the food group that the target crop falls in, i.e. they consume more of the crop in different forms. This is fairly straightforward – assuming that the diversity in production translates directly into diversity in consumption but in different processed forms.

The nutrition group at FAO has developed a very simple instrument for measuring dietary diversity. It is a yes/no response on whether food from each of nine groups has been eaten in the last day. Obviously it is very important when you do the interview and with whom you do it with. The nutrition group suggests administering the questionnaire to women of child-bearing age as they are the most likely to have eating patterns affected by income levels.

Measures of dietary diversity are optional – each case study group can decide if they want to include it. In order to include the dietary diversity instrument into the household survey it

will be necessary to modify the instrument to be relevant to the specific foods eaten in the survey site, and also to add some questions about the dietary diversity WITHIN the food category that the target crop falls into – in order to be able to test hypotheses (B). This can be done with assistance from FAO.

Resilience to shocks: Resilience to shocks is a measurement of the household's ability to maintain well-being in the face of adverse events – such as production failures, illness, or market failures. Again with this indicator of welfare, two levels of hypotheses can be tested: one specific to the target crop and production and the second more general.

A) Households with higher levels of infra-specific diversity within the target crop are less likely to experience a reduction in production of the crop due to pest, disease, drought or other causes of production failure. Here infra-specific diversity may be defined as number of varieties or the use of traditional vs. improved varieties.

B) Households with higher levels of crop diversity (interspecific together with infraspecific in the target crop) are more likely to be food secure. This is based on perceiving crop diversity as a portfolio decision on the part of farmers to manage risks and thus higher diversification would be associated with lower exposure to risks. However we may very well find the opposite because diversification out of agriculture into non-farm activities may be a much better way of managing co-variant risks.

This measure is included in the core set of variables to be collected across all studies and we have provided questions on the household template to determine it.

Wealth: We need a measure of wealth to condition behaviour and the management of the crop and variety portfolio. If we are concerned with the private good aspect of the risk-return tradeoff, then wealth is often the final state of concern although bounded rationality has questioned this assumption.

Income: To capture the household's agricultural portfolio management decisions we also must be concerned with the way in which farm resources are allocated with particular attention to labor, as it can be allocated to farm activities for wage income, non-agricultural on-farm labor earning activities. Furthermore, off-farm remittances can be an important determinant of behaviour if they are a buffer against downside risk.

ii. Utilization

The household surveys will be used to collect information on the size and scope of the household crop and variety portfolio and the factors that influence the management of this portfolio. The purpose is to answer questions about the private-good private-use value of the genetic services provided by seeds, to use this information to better understand local- and global-public good issues and to formulate better hypotheses at this level.

Portfolio Choice and Management: Portfolio management is often thought of in terms of the risk versus return tradeoff. In order to understand the portfolio management decisions that translate into land allocation decisions between and within crops, there is need to collect data on the sources of crop production risk (including variety-specific differences for target crops), price risk and the relative levels and covariate risk between crops and varieties. This requires price information over different states of nature in addition to production information.

A. Crop inventory

In order to address the portfolio management question, information about the area devoted to all crops grown on the farm is needed in addition to the more detailed target crop specific information.

B. Questions specific to the target crop

It is important to have some parallel information available for target crop varieties and other crops that make up the production (and implied consumption) portfolio. This includes information on variety turnover, traits supplied, historical usage, seeds source among others.

C. Triangular yield distributions

The use of these distributions is discussed by Hardaker et. al (2004). There is a theoretical basis for it (an approximation to the normal distribution) and moments can be derived, as well as marginal yield distributions and unconditional expected yields. Monte Carlo simulation can be applied to generate the distribution of the distribution. Using expected yields rather than observed in the previous season removes year-specific effects of yields, and SOME of the endogeneity (there are farmer specific effects, and we have tried to work these out in an instrumental variable approach).

It is not too difficult to implement as long as you use stones or straws to demonstrate to farmers. They tend to behave as if they are answering the schoolmasters' questions. It is sometimes burdensome, which is why we are suggesting that we focus on only the major, observable stresses, or group all biotic ones together.

There is measurement error in the denominator (plot area, though we will be using GIS in Mali, India, and Kenya and that will help), and measurement error in the numerator because farmers measure in terms of baskets, storage structures. In that case, one needs to measure the structures.

iii. Household Participation in Markets

We wish to evaluate the relationship between markets and crop diversity management so questions such as whether or not a household participates in the market must be answered and then second-level questions about the reasons discussed Are there barriers to participation in the input market? Output market? Is information and the manner in which price expectations are formed driving the participation and hence management decision? There are several factors in these questions that may devolve into, but are not limited to:

A. *Neoclassical price expectation and marketing margins*: This includes final product price issues as well as physical factors (transportation, stockage) affecting net returns.

*B. New Institutional Economics: Aspect*s of transaction costs including search, negotiation, bargaining, enforcement and policing aspects that control the actual transaction and not the price.

C. Social capital: The network of social relations to which household members are engaged can be an important determinant of seed sourcing decisions and have important impacts on information transmission.

D. Seed sourcing and market participation

These questions are derived from years of work at CIMMYT on maize seed management (Louette, Berthaud, Bellon, Aguirre, Smale), and built on other work done on other crops, by Heisey and Brennan, Byerlee, Morris, Almekinders and others who have used seed renewal concepts. They found that for both a) genetic diversity analysis and b) understanding farmer seed management for that crop, the unit of analysis needs to be the seed lot: the physical unit a farmer uses to reproduce a variety. A seed lot is plot, variety, and farmer-specific.

This approach was applied also in the case of minor millets by Nagarajan(2004), including self-pollinating and cross-pollinating species, and farmers' cultivars through improved varieties and hybrids. The concept is fundamental to differentiate between seed replacement (seed demand, relevant for the industry) and variety change (germplasm).

Why do we need this at the farm level? Here are several reasons:

1. We have argued that to understand market demand, you also need to understand non-market demand and management. The policy reason why is because we want to move the non-market demand into the market arena—address market failures by strengthening the seed system. Tripp (1998) and Sperling (2008), among others, emphasize continually on the need to understand both to be able to make progress, especially for minor crops.
2. Another reason is that we do not always have variety integrity in the market. We have product differentiation but not genetic differentiation.
3. At the farm level, you often have mixing, which affects genetic purity.
4. In Kenya, Mali and India, the seed replacement ratio at the farm level will be used as an impact indicator for access.
5. A variant of this was used in the case of bananas. For bananas, clones were used. Clearly clones are not mixed, but they are renewed because of disease and pest problems in the planting material. This type of work has been used in potatoes and tubers as well, for seed replacement and renewal.

iv. Household demographic variables

These are variables that are needed to control for inter-farm and market zone heterogeneity, including things such as household size, age and gender of head of household, ethnicity etc.

4. Hypotheses and Model

With some modification (unique to each of the case studies), we will be able to test some general hypotheses. With genetic characterization of samples drawn from the market and the field, in the context of a treatment sample design, it is possible to test the following hypothesis:

Hypothesis 1: Do market diversity levels differ between areas where interventions have occurred and those where they have not occurred?

Hypothesis 2: Do on farm diversity levels differ between areas where interventions have occurred and those where they have not occurred?

With a regression model, in a simultaneous (IV estimation) framework, or a "difference-in-difference" approach, , we will test:

Hypothesis 3: What are the determinants of a household's decision to participate in markets for seed? Is there a positive association between access to diversity in markets and participation in the input market?

Hypothesis 4: Controlling for the factors that determine participation, and for the level of diversity in markets, what is the impact of market participation on farm level diversity and farmer welfare?

In Mali, Kenya and India, farm level diversity will be measured with characterization from growouts on-station. Welfare will be measured in millet consumption and dietary diversity. Management of genetic resources will be measured with growouts, access to genetic resources will be will be measured using the market level information on variety availability, prices and information.

In order to evaluate the relationship between markets and agrobiodiversity, a number of empirical approaches may be used including matching, double-difference, and instrumental variable approaches[5]. Matching is used to enhance the suitability of the sample while other techniques are used to derive causal relationships. Algebraically, the econometric model for the instrumental approach is:

$$P_i = e + fZ_i + \upsilon_i$$

$$\omega_i = a + bP_i + cX_i + dM + \varepsilon_i$$

P = participation in the market.
ω= A set of impact indicators
Z = all of the variables that influence participation (including those found in X), as well as those that are correlated with neither ω nor the error term v
X = all of the variables that influence ω

[5] The reflexive methods might also be useful for stakeholders in their self-assessment, but without the application of economics methods.

M= fixed effects of market diversity, exogenous to the individual household.

In these econometric approaches, P is always a categorical variable (yes-no) (or with sufficient observations in each case multinomial). In this research, it will be possible to define the types of market participation:

1. autarkic
2. input market net buyer, output market net seller
3. input market net buyer, output market net buyer
4. input market net seller, output market net seller
5. input market net seller, output market net buyer
6. input market only – net buyer
7. input market only – net seller

With these market participation typologies defined, the relationship between market activities, transaction costs and other factors can be correlated with diversity variables and welfare indicators. The appropriate variables will depend on the context, but in principle, one of several hypotheses could be tested (several models can be run).

Again, some of these impact indicators are common across case studies, and some are unique to each case study. In Mali, for example, we wanted to look at seed replacement ratios, variety change, diversity of traits and attributes of varieties in stock. In some cases, wealth and income indicators could be tested. In several, production vulnerability (e.g. resilience) indicators are of interest. In most, consumption and dietary indicators are considered to be important.

If panel data is collected, a double-difference approach can be used to test for change over time. This model is formulated as:

$$\omega_{it} = aP_i + bt + cPt + dX_{it} + eM + \varepsilon_{it}$$

where

t = time period (beginning=0, end=1)
c = double difference effect
d = marginal effect of all of the variables that influence the heterogeneity of impact on management strategies (e.g., capital, human, social, physical endowments)
M=fixed effect, market diversity factors

Annex I: Measuring diversity in markets and on-farm:

A check list of minimum data set and calculations for variety diversity on-farm and in markets

The data set:

For each farm surveyed

1. Number of varieties grown
2. Area grown for each variety
3. Identity of each variety (name and whether traditional or modern)
4. Source of seed of each variety

For each community/village

5. Total area devoted to crop in the community
   n.b. not area sampled but total area of the community from which sampled farmers were drawn
6. Number of varieties
7. Identity of each variety (name and whether traditional or modern)

For each seller

8. Number of varieties sold
9. Amounts of each variety sold
10. Amounts of each variety available
11. Identity of each variety (name and whether traditional or modern)
12. Source of seed of each variety

For each market

13. Total amount of seed of crop sold in market
14. Total amount of seed of crop available in market
15. Identity of each variety (name and whether traditional or modern)

The calculations

Farm and community

1. Average number of varieties per farm (Richness, farm)
2. Average evenness (dominance, Simpson Index - SI) over farms
3. Total number of varieties in community (Richness, community)
4. Evenness (dominance, Simpson Index - SI) at community level,
5. Divergence among farms: Between farm SI/ Community SI (%)
6. Standard errors

Seller and market

1. Average number of varieties per seller (Richness, seller)
2. Average evenness (SI) of varieties sold over sellers
3. Average evenness (SI) of varieties available over sellers
4. Total number of varieties in market (Richness, market)
5. Evenness (SI) of varieties sold in markets
6. Evenness (SI) of varieties available in markets
7. Divergence among sellers: Between seller SI/ Market SI for varieties sold and varieties available.
8. Standard errors

Annex II: Socio-economic data that is needed at the market level

A check list of minimum data set and calculations for developing market level measures of access to diversity in the market (information flows, other types of transactions costs and prices)

The data set:

A. For each retail market selected for sampling[6]:
Instruments: market observation & key informant surveys

16. GPS coordinates (Lat and Long)
17. Frequency of operation
18. Total physical size (in area and number of stalls/vendors)
19. Presence and quality of market infrastructure
20. Transactions costs of selling via the market outlet (permits required, cost to obtain, storage requirements)
21. Rough estimate of composition of products sold at market (distinguished by retail outlet)
    5.a. list of all crops for which seed or product is sold in the market.
22. Characteristics of participant vendor(s) (farmer selling surplus, local vs. non-local traders, specialized seed distributors etc.)
23. For focus crop only:
    a. Grain (or output product) distinguished from seed in sales?
    b. Seed sales – improved or local varieties?
    c. Ways in which seed and varieties are distinguished (or not)
    d. Characteristics of the seed/grain that are demanded
    e. Physical unit of measurement used in sales

B. For each vendor
Instrument: Vendor survey

1. Characteristics of the vendor (gender, ethnicity, age, mother tongue, immigrant status)
2. Length of time selling in the market, distance to get to market, % income from sales at market
3. Transactions costs of selling in the market (permit/inspection fees/effort required, own or rent sales site)
4. Focus on crop questions: differentiate between seed & grain, seed sourcing, seed sorting & storage, factors determining price setting and changes in prices
5. Focus on crop seed lot characteristics: number of seed lots/varieties sold, number of years variety is sold, origin of variety, source of information about the seed lot, perception as improved or local, average amounts sold over recent past, factors determining price setting, average pricing over recent past, source of the seedlot and reason for source selection, price paid for the seedlot, information obtained from seller of the seedlot, information provided to the buyer of the seedlot.

The calculations (Need to be discussed/confirmed by group)

7. Market level measures of information availability about seeds and seed-lots (these will be rankings using aggregated vendor data on degree to which seeds distinguished from grain, the degree to which varieties are separated or mixed and information provided to buyers on seedlot sales)

[6] Note retail markets selected for sampling may take more than one form. They could be a weekly open-air market or they could be agro-veterinarian shops or other forms of market outlets. Some studies will have more than one type of market outlet identified as part of the retail market that will be sampled in these surveys.

8. Market level: average price over the selling season; measures of price stability and uniformity over the seed selling season;
9. Market level of seed vendor transactions costs (we could try to use absolute values and calculate mean, max and min but experience from other projects indicates that these data are hard to collect and not reliable – may be better to go for a ranking scale).

Appendix I: Templates for the key informant, market observation and vendor surveys

KEY INFORMANT Interview (The key informant may be either a vendor or other person familiar with the market characteristics and functions. This is a one time interview done prior to the period seed of the focus crop may be sold at the market – may be considered part of the value chain analysis).

(+ indicates question appears on other instruments; blue indicates answers from key informant interview are cross listed to market and or vendor instrument)

I. IDENTIFICATION

+Market Location **(Name of Village/Town)**	**MIDEN01A**	
+Market Name	**MIDEN01B**	
Date and time of interview	**MIDEN01C**	
Primary occupation of the key informant **Please describe**		
Primary clients of the key informant **Please list from most regular to occassional**		
How was the key informant recruited? **Please describe**		
+Is there an agriculture extension office in this village/town? (**1=Yes, 0=No**)	**MIDEN03**	
+Distance to nearest other comparable seed market **(Distance, Village/Town name)**	**MIDEN 04A**	**MIDEN04B**

II. CHARACTERISTICS OF THE "MARKET"

+Frequency of operation **(Days/Week)**	**MIDEN03**
+Hours of operation **(Hours/Day)**	**MIDEN03**
+In operation all year? **1=Yes, 0=No (if no, list when it is open)**	**MIDEN03**
Over what period is grain sold? (for case studies where focus crop is a grain) Over what period is seed sold? ? **Weeks/months**	
Are most of the seed vendors male or female?	
+Reliability (how often does market take place during scheduled time) **0=Never, 1=Rarely, 2=Sometimes, 4=Often 5=Always**	**MIDEN03**

+Reliability (how often does market take place during scheduled time) **0=Never, 1=Rarely, 2=Sometimes, 4=Often 5=Always**	**MIDEN03**
How long has the market been here?	
Does the number of vendors in the market change over the seed selling season? By how much? How much does the number of vendors in the market from year to year? What does it depend on?	
+Types of Vendors **LIST A** **Enter and code the different categories of vendors selling seed to farmers** **Circle the category this key informant falls into, if any Eg.:**	1. farmers selling surplus 2. local traders 3. non local traders 4. specialized seed distributors (wholesalers) 5. specific seed company representatives 6. other (please specify)
Types of clients/buyers of seed **LIST B**	Example: 1=Local farmers 2=Retailers 3=Government 4=Cooperatives 5= political candidates 6=Other – specify
Sources of seed for vendors **LIST C**	Example: 0=own fields 1=Local farmers 2=Retailers 3=Government 4=Cooperatives 5= political candidates 6=Other – specify

III. FOCUS CROP CHARACTERISTICS (MAXIMUM 2 FOCUS CROPS)

	CROP A	CROP B
+Type of varieties sold in this market (Example: 1=Creole, 2=open-pollinated improved variety (OPV), 3=Hybrid, 4=creolized 5=recycled hybrid) **LIST E**	**MCROPA01** **MCROPA02** **MCROPA03**	**MCROPB01** **MCROPB02** **MCROPB03**

+Are variety names recognized by vendors and purchasers? 1=Yes, 0=No			
+If yes, Named varieties available (Names) **LIST F**	**MCROPAv102** **MCROPAv202** **MCROPAv302** **MCROPAv402**	**MCROPBv102** **MCROPBv202** **MCROPBv302** **MCROPBv402**	
How many vendors sell these same varieties/types of seed?	**MCROPAv102** **MCROPAv202** **MCROPAv302** **MCROPAv402**		

Are there one or two vendors who dominant the sales of seeds for this crop in this market? If yes, what percentage share of total seed sales does each represent?		
+What is the peak sales period for each seed lot in a normal year? (beginning and ending weeks)		
+What are the low sales periods of the year? (beginning and ending weeks)		
What is the peak season in a bad year?		
+What are the peak sales hours during a given day?		
How do sellers decide what price to sell seeds/grain at? (does one seller have market power – sets the price or are prices regulated or is it mostly bargaining..etc.)		

IV. INSTITUTIONAL CHARACTERISTICS

Do vendors require permits to sell seed? 1=Yes, 0=No If yes, which category of vendors are required to have permits? How much do permits cost and how are they charged? (e.g. monthly, daily or volume)	**MINST01**
Are there other fees such as inspection fees that vendors have to pay? What are they and how much do they cost?	**MINST02**
Are there regulations on seed quality that are enforced in the market? 1=Yes, 0=No If yes, how? Market inspectors Labelling Other?	**MINST03**
Are there regulations on seed labeling enforced in the market? 1=Yes, 0=No If yes, how?	**MINST04**
Is information regarding seed quality regularly provided in the market? 1=Yes, 0=No If yes, how?	**MINST06**
+Does the market have requirements for storage of seed and grains? 1=Yes, 0=No If yes, are they generally complied with?	**MINST09**

+Does the market have requirements on sanitation 1=Yes, 0=No If yes, are they generally complied with?	**MINST10**
+Other regulatory presence 1=Yes, please specify 0=No	**MINST11**

VENDOR Survey: census or representative sample:
(+ indicates question appears on both key informant and vendor instrument; blue indicates answers from key informant interview are cross listed to market and or vendor instrument)

The vendors selected to answer this survey should be sellers of seed for the focus crop. In cases where there are relatively few vendors of seed for the focus crop in a market, the survey should be done for all (e.g. a census). In cases where there are many sellers (perhaps more than 10 we will need to develop a protocol for selecting vendors for interviewing. Data in the survey is collected at 3 scales: the vendor, the focus crop, and the seed lot of the focus crop. Some of the questions on the vendor survey (vendor and focus crop level questions) will only be asked once – these can be done in a first, initial visit or if they take too long they can be split between 2 visits.

We may want to ask questions on seed sales and prices on a regular basis (minimum once a week) over the period that seeds of the focus crop are sold on the market) This is something we need discussion on from the group.

I. VENDOR IDENTIFICATION

Name **VIDEN01** ID Number (assigned) **VIDEN02**	
Market Name	**VIDEN03**
Date and Time of Survey Start time: _____ Finish time: Number of purchases during interview_____- Typical quantity sold per purchase.....	**VIDEN03**
+LIST A - Type of Vendor Enter categories of vendors selling seed to farmers from key informant e.g.:	**1. farmer selling surplus** **2. local trader** **3. non local trader** **4. specialized seed distributor (wholesalers)** **5. specific seed company representative** **6. other (please specify)**

II. VENDOR CHARACTERISTICS

Age (Years)	**VIDEN04**
Gender: (1=Male, 2=Female)	**VIDEN05**
Mother Language (1=Official National Language, 2 =Dialect A, 3=Dialect B, 4=Other)	**VIDEN06**
Knowledge of Reading/Writing (0=No, 1=Yes)	**VIDEN07**
Education (Last grade completed)	**VIDEN08**
Ethnic group (categories should be decided as relevant in each case)	
From the community or immigrant (1=Community, 2=Immigrant, 3=Other)	**VIDEN10**
How long have you lived in the community? (Immigrants only)	**VIDEN10**
What do you consider your principal occupation (please list)	**VIDEN12**
Other occupations (List)	**VIDEN13**
How far did you travel to the market? (KM)	**VIDEN14**
How often do you sell at the market? 1. hours per day 2. days per week 3. weeks per year)	 **VIDEN15** **VIDEN15** **VIDEN15**
What percentage of your total income comes from seed sales (across all market/trades)?	
How long have you been selling at the market? (Years)	**VIDEN16**
Do you regularly sell at other markets? (0=No, 1=Yes) If yes, please list them:	**VIDEN17** 1. 2. 3.
Do you need a permit to sell in this market? (0=No 1=Yes)	
If so, how much does the permit cost? Do you pay once or daily?	**VIDEN28**
How many hours of effort did it take to apply for one? Hours of effort	**VIDEN29**
How many days did it take to receive one? days of waiting	**VIDEN27**
How long is the permit valid? indicate days or years	**VIDEN30**

Is it difficult to obtain a permit? (1=Easy, 2=Fairly Easy, 3=Fairly Difficult, 4=Difficult)	**VIDEN31**
Do you own or rent your stall?	
How much does it cost to own/rent a stall?	**VIDEN32**
Do you have a contract/agreement with the market **(0=No, 1=Yes)**	**PROD01**
If yes, please describe what the contract/agreement involves	
How often have you been inspected? If yes, did you have to pay a fine?	
III. QUESTIONS ON FOCUS CROP ONLY	
Do you differentiate grain from seed in your sales?	
To whom do you typically sell your seed?	**Proportion sold to 1=Local farmers 2=Retailers 3=Government 4=Cooperatives 5= political candidates 6=Other – specify** **VIDEN18** **VIDEN18** **VIDEN18** **VIDEN18 VIDEN18** **VIDEN18**
Do you have more than one source you obtain seed for this crop from?	
Do you buy the same type/variety of seed from multiple sources in one season? (0=No, 1=Yes)	**VIDEN22**
Do you sort and sell seed based on quality levels? (0=No, 1=Yes)	**VIDEN23**
Do you store your grain/seed? If yes, how do you store it? (Type of facility) Do you store grain separately from seed?	
What is your total grain/seed storage capacity?	

How much of your storage capacity do you utilize? Note to Patrick and Latha- we would like to get a measure of the total stock the vendor has available by estimating the total volumes that they store over the relevant season. I'm not sure the best way to ask this question – it would be useful if you could try out various means and give us feedback.	
Do you change your selling price for improved varieties based on what other vendors are charging? (0=No, 1=Yes) Do you change your selling price for farmer's varieties based on what other vendors are charging? (0=No, 1=Yes)	
If more people wanted your seed, would you raise your selling price? (0=No, 1=Yes)	
If your costs went up, would you raise your selling price? (0=No, 1=Yes)	

If fewer vendors were selling, would you raise your selling price? (0=No, 1=Yes)	
Does your selling price change depending on where you are selling? (0=No, 1=Yes)	
Does your price change according to whether it is a good year or bad year? (0=No, 1=Yes)	

IV.PRODUCT CHARACTERISTICS

Number of crops sold (total)		
+Enter LIST D – crops Characteristics of different crops sold	List types sold (Eg.modern, local) May use LIST E	Variety name if known May use LIST F
Crop A	**VPROD03A VPROD03B VPROD03C**	
Crop B	**VPROD04A VPROD04B VPROD04C**	
Crop C	**VPROD05A VPROD05B VPROD05C**	
Crop D	**VPROD06A VPROD06B VPROD06C**	
Crop E	**VPROD07A VPROD07B VPROD07C**	
Crop F	**VPROD08A VPROD08B VPROD08C**	
Crop G	**VPROD09A VPROD09B VPROD09C**	
Crop H	**VPROD10A VPROD010B VPROD010C**	

V. FOCUS CROP SEED LOT CHARACTERISTICS

For focus crop:	**SEED LOT A**	**SEED LOT B**	**SEED LOT C**	**SEED LOT D**
Name of variety or seed lot	**VCROP01A**	**VCROP01B**	**VCROP01C**	**VCROP01D**
For how many years have you sold this variety?	**VCROPAv102 VCROPAv202 VCROPAv302 VCROPAv402**	**VCROPBv102 VCROPBv202 VCROPBv302 VCROPBv402**	**VCROPCv102 VCROPCv202 VCROPCv302 VCROPCv402**	**VCROPDv102 VCROPDv202 VCROPDv302 VCROPDv402**

How did you originally find out about the variety? (multiple answers) 1=Saw in the field 2=Farmer requested 3=Technician/Extension recommended 4=Another vendor recommended 5=Source recommended 6=Saw it in a demonstration plot 7=Saw it in a brochure 8=Other (Specify)	**VCROPAv127**	**VCROPBv127**	**VCROPCv127**	**VCROPDv127**
What is its region/village of origin? (for landraces only)				
How do you consider this seed type: Enter LIST E – seed type, eg 1=Farmer variety 2=Open-pollinated improved variety (OPV), 3= hybrid (each case study and crop should set categories s relevant)	**VCROPAv120**	**VCROPBv120**	**VCROPCv120**	**VCROPDv120**
How much do you sell of this variety over a normal season?				
Do you sell more or less of this variety in a bad year? 0=more, 1=less				
*Amount sold on an average day mid-season specify units	**VCROPAv104**	**VCROPBv104**	**VCROPCv104**	**VCROPDv104**
*Over the past 3 years, what's the maximum amount of this seedlot you've sold over a season? specify units				
*Over the past 3 years, what's the minimum amount of this seedlot you've sold over a season? specify units				

Approximate total amount sold last season/year at this market? Specify units	**VCROPAv104**	**VCROPBv104**	**VCROPCv104**	**VCROPDv104**
Do you have a fixed price or do you negotiate?				
Do you have different kinds of sales agreements for selling seeds to farmers? (0=No, 1=Yes) If yes, what types of arrangements do you have? a) cash sales b) sales on credit c) discounted sales for bulk purchases d) discounted sales for preferred customers				
*Over the past 3 years what was the average price you've sold this seedlot for?				
*Over the past 3 years what was the lowest price you sold it for?				
*Over the past 3 years what was the highest price you sold it for?				
What characteristics of the seed lot made you decide to stock it? (multiple answers) 1= Many farmers request it 2= Farmers will pay a good price for it 3= Credit obtainable from seed source 4=its easy to get locally.. (specify)				
From whom did you procure this seed lot Enter LIST C - sources Eg. (0=own farm 1=farmer 2=local trader 3=nonlocal trader 4=wholesaler 5=seed co. 6=other, specify)	**VCROPAv121**	**VCROPBv121**	**VCROPCv121**	**VCROPDv121**
Do you personally know the source? (0=No, 1=Yes)	**VCROPAv121**	**VCROPBv121**	**VCROPCv121**	**VCROPDv121**
What do you pay for the seed? (price/unit) What are the terms of the sale?	**VCROPAv108**	**VCROPBv108**	**VCROPCv108**	**VCROPDv108**

Cash payment Credit (terms) in kind?				
Where is this seed lot source located? Place, Municipality	**VCROPAv122**	**VCROPBv122**	**VCROPCv122**	**VCROPDv122**
Distance from the seed source (Km.)	**VCROPAv125**	**VCROPBv125**	**VCROPCv125**	**VCROPDv125**
Why did you choose this source to obtain seed? (May choose more than one answer, as many as necessary) 1=Only source known with desired variety, 2=Cheaper/est source, 3=easiet to obtain, 4=most trusted sourc 5= had obtained it there before 6=Other (specify) they extend credit; part of a family network	**VCROPAv123**	**VCROPBv123**	**VCROPCv123**	**VCROPDv123**
What information about the seed lot did the source provide you with? a) variety name b) origin of the seed c) production traits (length to maturity, tolerance to stresses) d) consumption traits				
What information about the seed lot do you provide to buyers? e) variety name f) origin of the seed g) production traits (length to maturity, tolerance to stresses) h) consumption traits				

	SEED LOT A	SEED LOT B	SEED LOT C	SEED LOT D
How much of this seed lot did you sell over the past week (or whatever relevant time frame?)				
Approximately how many buyers did you have?				
Did all buyers pay the same price? (for the same seed lot) If no, why not? a) preferred custormer, credit arrangement, bulk sales... Did any receive credit?				
What was the highest price this seedlot sold for this week? What was the lowest? Local currency/unit (bag, etc.)	**VCROPAv105** **VCROPAv205** **VCROPAv305** **VCROPAv405** **VCROPAv505** **VCROPAv605**	**VCROPBv105** **VCROPBv205** **VCROPBv305** **VCROPBv405** **VCROPBv505** **VCROPBv605**	**VCROPCv105** **VCROPCv205** **VCROPCv305** **VCROPCv405** **VCROPCv505** **VCROPCv605**	**VCROPDv105** **VCROPDv205** **VCROPDv305** **VCROPDv405** **VCROPDv505** **VCROPDv605**

MARKET Observations (conducted by a member of the research team, data collected via observation and questions to participants in the market, survey is conducted one time only prior to period for which seed of the focus crop may be on the market)

(+ indicates question appears on both key informant and market instrument; green indicates answers from key informant interview are cross listed to market and or vendor instrument)

I. VENDOR IDENTIFICATION

+Market Location **(Name of Village/Town)** + Market Name Date and time of observations **Market area in square meters** **GPS Coordinates (Latitude and Longitud)**	**MIDEN01A** **MIDEN01B** **MIDEN01C**	
	MIDEN02	
+ Is there an agriculture extension office in this village/town? (**1=Yes, 0=No**)	**MIDEN03**	
+ Distance to nearest other comparable seed market **(Distance, Village/Town name)**	**MIDEN 04A**	**MIDEN04B**

**Attach available DRAWINGS, MAPS, PHOTOS OF MARKET AND MARKET SITUATION

II. PHYSICAL CHARACTERISTICS OF THE "MARKET"

Market area **(Acres)**	**MMRKT00**

Permanent	**MMRKT03A1**	**MMRKT03B1**
Semi-permanent		
Makeshift/Temporary	**MMRKT03A2**	**MMRKT03B2**
Covering on stalls (Quantity of stalls / Quality)		
Covered	**MMRKT03A3**	**MMRKT03B3**
Uncovered	**MMRKT03A4**	**MMRKT03B4**
Type of Infrastructure (Quantity / Quality)	**Quantity**	**Quality (1= Good, 2=Medium, 3=Poor)**
Public Toilets	**MMRKT01A1**	**MMRKT01B1**
Public Sinks	**MMRKT01A2**	**MMRKT01B2**
Public Covered Areas	**MMRKT01A3**	**MMRKT01B3**
Road access in and out	**MMRKT01A4**	**MMRKT01B4**
Lighting	**MMRKT01A5**	**MMRKT01B5**
Drainage	**MMRKT01A6**	**MMRKT01B6**
Quality of pest management	**(1= Good, 2=Medium, 3=Poor)**	
Rats	**MMRKT10A1**	**MMRKT10B1**
Insects in market	**MMRKT10A2**	**MMRKT10B2**
Quality of sanitation	**(1= Good, 2=Medium, 3=Poor)**	
Water (sink, open spigot)	**MMRKT11A**	
Toilets	**MMRKT11B**	
Cleanliness of tables on which goods are being sold	**MMRKT11C**	
Other (specify)	**MMRKT11D**	

III. CHARACTERISTICS OF ALL SEED SOURCES

A. – PARTICIPANT CHARACTERISTICS		
Number of stalls selling seeds/planting materials **(Number)**	**MMRKT15**	
+Enter/cross-check LIST A - types of vendors selling seed to farmers below, eg.:	Number of vendors	Is credit available to farmers: (yes = 1, no =0)
1. farmers selling surplus	**MMRKT1**	**MMRKT1**
2. local traders	**MMRKT1**	**MMRKT1**
3. non local traders	**MMRKT1**	**MMRKT1**
4. specialized seed distributors (wholesalers)	**MMRKT1**	**MMRKT1**
5. specific seed company representatives	**MMRKT1**	**MMRKT1**
6. other (please specify)	**MMRKT1**	**MMRKT1**
B. – PRODUCT CHARACTERISTICS (SEE MALI SURVEY FOR EXAMPLES OF HOW THIS WAS LAID OUT)		
Availability of other products at the market	**1=Yes, 0=No**	
Fertilizer	**MMRKT42A1**	
Pesticide	**MMRKT42A2**	
Cooking oil	**MMRKT42A3**	
Household items (brooms, cleaning items, etc.)	**MMRKT42A4**	
Clothing	**MMRKT42A5**	
Other (specify)	**MMRKT42A7**	

+Enter/cross check LIST D - types/names of crops sold	Number of vendors	Approx number of named varieties if posted	Seed (0), Grain (1), Mixed Pool (3)
Crop A	**MMRKT43B**	**MMRKT43C**	**MMRKT43D**
Crop B	**MMRKT44B**	**MMRKT44C**	**MMRKT44D**
Crop C	**MMRKT45B**	**MMRKT45C**	**MMRKT45D**
Crop D	**MMRKT46B**	**MMRKT46C**	**MMRKT46D**
Crop E	**MMRKT47B**	**MMRKT47C**	**MMRKT47D**
Crop F	**MMRKT48B**	**MMRKT48C**	**MMRKT48D**
Crop G	**MMRKT49B**	**MMRKT49C**	**MMRKT49D**
Crop H	**MMRKT50B**	**MMRKT50C**	**MMRKT50D**
Crop I	**MMRKT51B**	**MMRKT51C**	**MMRKT51D**
Crop J	**MMRKT52B**	**MMRKT52C**	**MMRKT52D**
Other (specify)	**MMRKT53B**	**MMRKT53C**	**MMRKT53D**

IV. FOCUS CROP CHARACTERISTICS (NUMBER SPECIFIC TO EACH CASE – THIS EG. IS 2 CROPS WITH WHATEVER NUMBER OF VARIETIES (THIS SECTION HAS BEEN REVISED TO USE THE MALI QUESTIONS (E.G. I COPIED THE ENTIRE SECTION FROM MELINDA'S REVISED VERSION)

	Characteristic	Crop 1	Crop 2	Crop 3
1	Sales of local varieties (yes=1; no=2)			
2	Sales of grain from harvest of improved seed (yes=1; no=2)			
3	Sales of new seed of improved varieties (yes=1; no=2)			
4	Do vendors recognize variety names for the grain they sell? (yes=1; no=2)			
5	If so, often or rarely? (often=1; rarely=2)			
6	If not, which characteristics do vendors use to distinguish one grain type from another? (list in the columns by crop)			
7	Do purchasers ask the names of the varieties for the grain they are considering? (yes=1 ; no=2)			
8	If so, often or rarely? (often=1; rarely=2)			
9	If not, do they ask about characteristics of the grain?			
10	If so, which characteristics? (list in the columns by crop)			
11	Is seed sold in the form of grain? (yes=1 ; no=2)			
12	Do vendors know when purchasers are asking for seed rather than grain? yes=1 ; no=2) Explain_____			
13	Do vendors mix varieties or types in the same container or do they separate them ? (separate=1; mix=2)			
14	Do vendors sort their types or varieties by quality? (yes =1 ; no=2)			
15	What is the physical unit of measurement for sales ? (sack, bowl, basket)			
16	What is the weight of the unit in kgs ?			
17	Do vendors stock their grain/seed?			
18	If so, where ? (list in the column, by crop)			

Appendix II: Templates for the household surveys (Instruments)

Household Survey Instruments	
1. Household member inventory, including non residents	Base
2. Household assets	Base
2.1 Agricultural implements	Base
2.2 Non-ag implements	Base
2.3 Household construction	Base
2.4 Livestock holdings	Base
2.5 Crops in storage	Base
3. Land holdings	Base
3.1 Owned land operated by farmer	Base
3.2 Rented land	Conditional
3.3 Tenure	Optional
3.4 Cropping system	Base
3.5 Production and utilization	Base
4. Target Crops	
4.1 Rate of Variety Change	Base
4.2 Variety Use and Sourcing - planted varieties	Base
4.3 Variety Use and Sourcing - stocked varieties	Base
4.4 Production under Stresses	Needs discussion - see sheet
4.5 Diversity of Variety Attributes	Base
4.6. Target Crop Questions for Non-producers	Conditional
5. Non-farm income	
5.1 Business-Informal	Optional - if expenditures used instead (table 8.1)
5.2 Salaried	Optional - if expenditures used instead (table 8.1)
6. Input and Output Market Activities of Target Crops	Base
6.1 Seed exchanges	Base
6.2 Seed market price formation/transactions	Base
6.3 Output market participation	Optional? perhaps should be dropped if too repetitive
6.4 Output market exchanges	Base
6.5 Output market price formation/transactions	Base
7. Organizational Affiliation and Role	Optional
8. Consumption	
8.1 30-day expenditure (can be modified to any period)	Optional - if income is used instead (tables 5.1 and 5.2)
8.2 Rapid consumption	Optional
8.3 Consumption resilience i.e. annual variation	Optional

REFERENCES:

Jayne T.S., J. Govereh, M. Wanzala, M. Demeke, (2003), **Fertilizer market development: a comparative analysis of Ethiopia, Kenya, and Zambia,** *Food Policy,* 28: 293-316

World Development Report, (2002), **Building Institutions for Markets**, World Bank, Oxford University Press, Washington DC, US.

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

Sadoulet, E. and A. de Janvry, (2003), **Quantitative Development Policy Analysis**, *The Johns Hopkins University Press*, Baltimore MD, US.

Bellemare, M. and C. Barrett, (2006), **An ordered Tobit Model of Market Participation: Evidence from Kenya and Ethiopia,** *American Journal of Agricultural Economics,* 88 (2): 324–337.

Dorward, A., J. Kydd, J. Morrison, C. Poulton, (2005), **Institutions, Markets and Economic Co-ordination: Linking Development Policy to Theory and Praxis,** *Development and Change,* 36 (1): 1–25

Morris Michael L., J. Rusike and M. Smale, chapter X, **"Maize Seed Industries: A Conceptual Framework"**, in Morris, Michael L. (1998), *Maize seed industries in developing countries*, London: Lynne Rienner

Badstue, Lone B., M. Bellon, X. Juárez , I. Manuel , and Ana M.Solano, (2002), **Social relations and seed transactions among small-scale maize farmers in the Central Valleys of Oaxaca, Mexico,** ESA Working Paper No. 04-16.
http://www.fao.org/es/ESA/en/pubs_nrm.htm

Lambert D. and Wilson W., (2003), **Valuing Varieties with Imperfect Output Quality Measurement,** *American Journal of Agricultural Economics,* 85 (1): 95–107

Barkley Andrew P. and L. Porter., (1996), **The Determinants of Wheat Variety Selection in Kansas, 1974 to 1993,** *American Journal of Agricultural Economics*, 78 (1): 202-211.

McCann, J.C.,(2005), **Maize and Grace: Africa's Encounter with a New World Crop, 1500–2000.** Cambridge: Harvard University Press. Pp. xiii, 289.

Ó' Gráda C., (1999), **Black '47 and beyond: The great Irish famine in history, economy, and memory**. Princeton, NJ: Princeton University Press.

Fafchamps M. and R. Vargas, (2005), **Selling at the Farmgate or Travelling to Market,** *American Journal of Agricultural Economics*, 87 (3): 717–734

Smale, Melinda, (2006), **Valuing Crop Biodiversity: On-farm Genetic Resources and Economic Change** , *CABI Publishing*, Wallingford, UK.

Hartl, D.L. & A.G. Clark, (1997), **Principles of Population Genetics**, Sinauer Associates, Sunderland, MA.

Hill M. O., (1973), **Diversity and Evenness: A Unifying Notation and Its Consequences**, *Ecology*, 54 (2): 427-432

Tothmeresz, B. (1995), **Comparison of Different Methods for Diversity Ordering,** *Journal of Vegetation Science*, 6(2): 283-290.

D. **Mouillot** and A. Leprêtre, (1999), **A comparison of species diversity estimators**, *Researches on Population Ecology,* 41: 203-215.

Nagarajan, L.and M.Smale, (2005), **"Local Seed Systems and Village Level Determinants of Millet Crop Diversity in Marginal Environments of India"**, Joint Discussion Paper of IFPRI/ICRISAT and FAO published by EPTD division, IFPRI. March, 2005.

Nagarajan, Latha, (2004), "**Managing Millet Diversity: Farmers' choices, seed systems and genetic resource policy in India"**. PhD dissertation, Department of Applied and Agricultural Economics, University of Minnesota, St.Paul, USA. December 2004.

Ruel, M. T.,(2003), "**Operationalizing Dietary Diversity: a Review of Measurement Issues and Research Priorities"**. *Journal of Nutrition.* 133: 3911S–3926S

Hardaker J. B., R.B.M. Huirne, J. Anderson, L. Gudbrand, (2004), "**Copying with Risk in Agriculture, 2nd Edition",** *CABI Publishing*, Wallingford, UK.

Anderson, J.R.; J.L. Dillon, J.B Hardaker, (1977), "**Agricultural decision analysis"**, Iowa State Univ. Press, Ames, Iowa (USA)

Tripp, R., (1998), **New Seeds Old Laws: Regulatory Reform and the Diversification of National Seed Systems.** London, Intermediate Technology Publications, pp 88-120.

Sperling L., H.D. Cooper and T. Remington, (2008), "**Moving towards more effective seed aid",** *Journal of Development Studies ,* Volume 44, Issue 4 : 586 – 612

Edmeades S., J. M. Jackson and M. Smale, Chapter 9, "**Use of Hybrid Cultivars in Kagera Region,Tanzania, and Their Impact",** in Smale M. and W. K.Tushemereirwe**,** (2007), An Economic Assessment of Banana Genetic Improvement and Innovation in the Lake Victoria Region of Uganda and Tanzania**,** *Research Report 155***,** International Food Policy Research Institute IFPRI, Washington D.C. US.

Ravallion, Martin., (2005), **"Evaluating Anti-Poverty Programs".** *Policy Research Working Paper 3625*, Washington, D.C.: World Bank.