A metapopulation approach to farmer seed systems: Methodology for agricultural biodiversity conservation policy

M. Eric Van Dusen
UC Berkeley

Keywords: Agricultural Biodiversity, Seed Systems,
Conservation Policy, Household-Farm models
JEL codes: O12,O17, Q12,Q2

Introduction

This work is being developed in order to provide methodology and empirical approaches for understanding farmer seed systems in developing countries. A meta-population is composed of a number of nodes of smaller populations inhabiting ecological niches connected by migration and colonization. A farmer seed system is composed of a series of farmers and the different ways in which genetic material is exchanged and moves between farmers, as well as the ways that genetic material is selected and shaped by each farmer's behavior. This paper has two components, first an outline of key points in integrating population modeling with farm level information, and second, an application with household survey data from Mexico.

A starting point for this approach is to improve the understanding of farmer seed systems, utilizing information on farmer management of seeds to move from the scale of an individual farmer to the level of a local population, composed of a region of farmers interacting through seed exchange. This population approach to seed system is a needed policy tool in three areas: 1) the *in situ* and on farm conservation of agricultural biodiversity, 2) modeling the possible impacts of the release of Genetically Modified (GM) crops in areas of high genetic diversity, and 3) improving access to genetic materials where institutional and environmental efforts have limited success in the past. In all of these areas there are important policy questions about how social and economic factors affect genetic material in farmers fields, and how market and institutional forces work at both local and regional scales.

The empirical application explores social, economic, and institutional factors that can influence the seed system. Statistical and econometric examples of how this approach can be applied are provided in order to illustrate the different possible components of a seed system model. Data from an original household survey in Puebla, Mexico designed to capture the variation in ecological and market conditions, are used to illustrate a seed system and test economic models. Statistical approaches are presented for geneflow, effective population size, seed source and seed replacement. Regression results are presented for the age of seed lots held by farmers with respect to agro-ecological, social and economic variables.

The idea of a farmer seed system is not particularly well defined but is receiving increasing attention from development organizations (ODI, UN-FAO, World Bank). A seed system is a group of concepts, on farmer selection, exchange of materials, and either extinction and loss of old materials or acquisition of new materials (Wright and Turner). One possible definition would be all informal channels that farmers use to access genetic materials, outside of the formal, commercial and private, seed sectors. The idea of modeling a farmers seed system as a crop metapopulation has been mentioned as a useful model (Louette, Berthaud, Brush et al) and suggested as an avenue for future research (IPGRI), but has not been formalized. Quantitative approaches to a farmer seed system have been proposed (Byerlee and Heisey) and different components have been estimated from survey or regional data, but the entire system has not been modeled. Although the definitions of these terms are evolving, this paper proposes the policy relevance for this approach, and illustrates how it could be applied.

Policy Relevance

In situ and on farm conservation of Crop Genetic Resources

There are three areas where the development of this new methodology and generation of empirical case studies are currently needed. This approach is first motivated by a need to evolve our approach to in situ conservation of agricultural biodiversity, also called Crop Genetic Resources (CGR). Current approaches used a household based approach to look at the social and market context of households choosing to conserve traditional varieties, merging tools of ethno-botany with household-farm modeling. There is a need for in situ studies to move to a community or regional level and to be able to look at the dynamic evolution of a crop population across a set of farmers.

Several authors have proposed that the dynamic behavior of farmers (with respect to seed management) is the most important focus for a policy on conservation. Farmers are constantly selecting materials, crops are continuously evolving, and if a policy is to target conservation, the target should be this evolutionary environment, where both crops and farmers are responding to ecological conditions. The metapopulation approach is a technique for modeling dynamic and evolutionary contexts, using information from cross-sectional surveys such recall and reference techniques to develop dynamic parameters.

Under commitments to the CBD and IT-PGRFA, each country should have a national plan for conservation of agricultural biodiversity. Despite ongoing debates on the relevance of on farm conservation, under the GEF, and under integrated conservation and development plans, funding is being allocated to on farm projects. The metapopulation methodology can contribute to the assessment of the variety of new approaches being implemented.

Biosafety

The second key area of policy relevance is in the realm of agricultural biosafety regulating the field release of genetically modified crops. While there is some research on the biological and ecological modeling of the effects of the release (or escape) of genetically modified crops into areas of related biological diversity, there is a concurrent need to investigate human aspects of the crop management that affect the spread or control of foreign genes in a farmer environment.

A highly visible example is the controversy surrounding the detection of transgenes in local varieties of maize in Oaxaca. While there has been a hotly contested debate concerning the accuracy of detection methods, there has been scant discussion of the fundamental social science questions such as how did it get there, what can be done to contain it, and how does it impact the livelihoods of the farmers? Several fundamental biosafety questions (such as how will transgene fare when escaped into a farmer population? or how far will it spread?) have economic and social science components for policy . Questions of selection pressures have an inherent human bias that needs information from field study: how do farmers select seed, what other selection pressures are applied to crop populations through growing conditions and management. A related question is how far does genetic material travel, which will be determined by combination of biological and human forces.

This information documenting farmer practices and modeling the effects on crop populations is necessary for a regional or national biosafety commission to make decisions. Examples include whether to release a specific gene in a specific crop in a given region, is a limited release to a limited area possible, and what options exist for containment, control, or mitigation. (Possibly the most expensive biosafety containment efforts to date in the US were the Starlink maize and Prodigene maize. Both costly errors were not due to a lack of biological or ecological information, but resulted from a lack of understanding and communication about farmer behavior and farm level conditions). National governments have a commitment under the Cartagena Protocol to review crops with a national biosafety commission, and the UNDP has been funding capacity building for the forming of these national commissions. Without local level information, and methodology for modeling the impacts of farmer behavior, the policy impacts of these commissions will be limited.

Access to CGR

The third area is to provide policies to facilitate access to CGR, whether through the dissemination of the products of crop breeding or identifying constraints to access in local seed systems. For many poor farmers, the formal seed supply systems of the green revolution and CGIAR model have had a limited impact, where farmers could gain from improved germplasm but solely source seed through informal networks. Previous Green Revolution studies focused on why farmers either failed to adopt new seed technologies, or how governments and seed companied failed to provide access. A better understanding of the informal networks, and policies adapted to their economic and institutional contexts, are necessary to extend the benefits of access to crop genetic resources.

Farmer to farmer exchange was fundamental to the rapid dissemination of materials during the expansion phase of the Green Revolution. In many areas of early Green Revolution successes further gains were limited by low turnover of varieties. The sustained growth in crop productivity depends on crop breeders continuing to improve materials and incorporate new genetic material, while simultaneously pests evolve and the plants' resistance to pests breaks down over time. However, it has also been documented that farmers stay on a sub-"optimal" path of keeping older varieties, and therefore the process to model is not a discrete technology adoption decision, but an adoption path with a different evolutionary framework (Heisey et al).

In many non-favored environments, farmer-to-farmer exchange of seed is the principal source of material, and policies that focus solely on the institutional environment for increasing the role of private seed companies can be misguided. Furthermore there is a growing popularity of participatory approaches to breeding and varietal selection. This is partially due to the perceived gains from combining breeding and extension to improve access and accelerate diffusion rates, taking into account farmer preferences and utilizing pre-existing farmer seed networks (Witcombe et al) . Finally there have been movements to change disaster seed relief programs to take into account farmer seed networks and source seed locally or regionally. These new approaches seek to strengthen robust farmer seed systems and perhaps improve the efficiency over importing seeds.

A farmer seed system approach is indeed a convergence of many different evolving ideas about how to improve access to CGR. National governments and international agencies have to develop practical and working solutions for Access and Benefit Sharing under CBD and the IT-PGR. An analytical and quantifiable approach to farmer seed systems approach will be necessary to both reconcile and integrate the apparantly diverging goals of conservation of diverse materials and facilitated access to improved materials.

Metapopulation theory

The idea of a meta-population comes from conservation biology, and the current work of a group of theoretical ecologists. While it is hard to apply to the crop case with the complexity used in the ecological literature, it is a useful framework for the mechanisms of farmer-based seed systems. The ecological modeling derives from studies of Island Bio-geography, where different species inhabit different ecological niches across a fragmented landscape, and there are movements of colonization and migration between niches and local trends of growth or extinctions within niches (MacAurthur and Wilson). An antecedent to metapopulations was the idea of mainland-island geography, where a spatially disaggregated population consists of a mainland that keeps sending colonizers out to island niches. The overall population is in all of the possible places, despite inhabiting distinct niches geographically. Based on parameters such as the distance between the mainland and the island, and the rate of migration, equilibrium conditions can be reached.

A metapopulation was a further level of abstraction, defined by a series of patches of similar ecological conditions (similar suitability as habitat), any number of which could be inhabited at any given time, driven by the rate of migration and the rate of extinction. Levins' model of a metapopulation was dynamic, where differential equations describe the probabilities of a patch being populated. This corresponded to an overall shift in ecology towards dynamic modeling, which can introduce a long-term perspective for conservation, but can limit the empirical application. Using the quantification of rates of extinction and migration, it is possible to derive mathematically a minimum number of patches to support a population or the minimum number of individuals within each patch.

Characteristics for empirically modeling metapopulations are an underlying environment with spatial heterogeneity, a structure of nodes and linkages between them, with population probabilities and migration rates. Some of the key results from the literature are:

- Distance and distribution of patches matters – physical characteristics of the spatial heterogeneity can drive results. The size and number of patches or the distance between them can impact the likelihood of successful migration.
- Success of Colonization – It is necessary to model both the probability of migration to a new patch, and the probabilities of successful establishment in the new patch to recover a realistic migration parameter.
- Successful establishment depends on multiple factors - Establishment in a new patch can depend on a new set of patch-specific factors, such as the intensity of competition and predation, or the probability of finding mates.
- Local extinction possible - As long as there is some degree of migration, local extinction in any given patch is possible. The patch may still be a vital part of the overall intertemporal habitat, but may not be inhabited at any given time.
- Extinction Debt – A population may be present in a patch but declining due to a low number of individuals or a limited genetic pool. A population may appear as present in a site sample, but may be below an extinction threshold and in a dynamic sense be in the process of disappearing.
- Genetic Rescue – A population can be declining (suffering from extinction debt) but just enough new individuals come (not total replacement) to rescue the viability of the local sub-population. The relevant migration rate for the restoration of a local population or inclusion of a patch in a healthy metapopulation, may be just the migration rate to get above the extinction threshold, not that for total colonization.

Crop metapopulations

In the crop case, the implication would be to move from each farmer having a distinct population to the population being composed of a community or set of farmers in an ecologic or agronomic region. In an informal seed system, the rates of colonization are the exchanges and sources of seeds of farmers within the set of farmers, and colonization and extinction are the turnover rate and adoption of new seed types by farmers. As the processes of seed exchange and varietal adoption have social and economic components, and are mediated by local institutions in farm communities, there is an opportunity to empirically look for the social and economic factors influencing the linkages between the patches in the crop metapopulation.

Some of the characteristics of a crop metapopulation would be:

- Individual farmers manage local populations, and are linked through seed exchange and gene flow.
- Patchiness – the agricultural environment, although assumed homogenous in other models, is very heterogeneous, varying within a given farm, a village, or a region. Farmers can identify the different agro-ecological conditions have ways to adjust their crops and agricultural technologies.
- Adaptation - seeds are adapted to local conditions, some serve specific niches, some perform across a range of agro-ecological conditions. Each seed type contains a bundle of genetic traits that has a set of genotype by environment interactions for any given growing conditions. Farmers can perceive this and seek genetic material suited to their conditions.
- Multiple seed traits – the constructed definitions of patchiness or adaptation can be extended to other seed traits. Beyond matching to traits such as soil type or annual rainfall, regional differentiation can cause farmers to seek adaptation to social or economic niches such as labor availability, consumption preferences, or differential risk markets.
- Farmers experience loss of seed (*local extinction*), but this is mitigated through seed exchange (*migration*). Farmers may lose their own saved seed in any given year (due to weather, pests, hunger, etc.) but will source similar seed from a neighbor and continue planting a given variety. Farmers may even stop planting a given variety any one year, but know they can resume in the future as long as they can still find a reliable seed source who has continued to save seed.

The relevance for policy analysis lies in determining what is the unit and scale for conservation, and what aspects of the overall system are necessary to conserve. Varieties with value for in situ conservation of agricultural biodiversity compete for the same land area with other uses and suffer from fragmentation of habitat. The conservation question lies in how many populations is the minimum number necessary for survival, and how many linkages between those populations need to be conserved. It is important to also note that biological and genetic characteristics of different crops will drive the results. Self pollinated crops like wheat and rice generally breed true to type over successive generations and suffer little from geneflow or pollen contamination. Open pollinated crops like maize have a higher level of diversity within each farmers population, are more affected by pollen flow, and genetic purity is more difficult to maintain. Vegetatively propagated crops like potatoes or bananas have lower levels of diversity because they are sourced through clones and not seeds, but the farmer and market networks are increasingly important, as well as the fact that it may become easier to maintain several distinct types.

Probability of Presence or Absence

In a crop metapopulation, whether a given habitat patch is inhabited would correspond to the probability that a given farmer is planting a given crop or variety. Each farmer with the agro-ecological and other conditions that correspond to a variety's traits can be considered as potential habitat for that population. On a larger scale each village or micro region in a larger population could be the unit of analysis. In either case the approach used has been a household-farm model of farmer activity choice. In different applications a household-farm model is extended to test different possible hypotheses and incorporate various constraints such as input fixities, incomplete markets, risk aversion, or a nesting of a range of multiple cases. This approach is relatively rich in modeling and empirical applications, from literature documenting the adoption of modern varieties, including studies exploring incomplete and partial adoption, and newer literature on on farm *in situ* conservation. The important outputs from these models are linkages between socio-economic variables and crop diversity outcomes. This approach has been advocated for policy outcomes such as identifying causes of de facto conservation, where farmers conserve without policy interventions, for identifying farmers with the lowest opportunity costs or identifying the forces of economic development that destabilize traditional crop populations.

Spatial Structure

The spatial structure of a crop metapopulation would be constructed from nodes – farmers or farm plots, and linkages – how farmers source seed from each other. A starting point for nodes would be all farm plots that have similar agro-ecological conditions. This could be extended to regional forces such as market or cultural factors that keep a particular variety or its consumption characteristics in demand. Nodes can also be defined by market conditions, such as the use of tractors, synthetic fertilizers, or herbicides, all of which can drive crop and variety choices and depend on spatial or regional location. Finally spatial distribution and fragmentation may need to take into account forces of competition, such as commercial crops or other higher value land use possibilities.

The linkages in the crop case are the interactions between farmers in order to source seeds. Within a given population of farmers a key parameter for the spatial structure is how often a member of the population will source seed from outside of the local network. Which farmers and how often they bring seeds in will drive the rates at which a system is open or closed. Within a given system it is also important to determine whether seed is sourced from a single source or a small group of "local experts" or is seed sourced uniformly from all possible members of a population.

Decisions about the basic unit of analysis are implicit in designing the structure of a crop metapopulation. From the literature on seed management the farmer is usually the basic decision making unit, the individual agent of conservation, selection or exchange, and the manager of a crop population. Alternatively it is possible to use each discrete plot managed by a farmer as the unit of analysis, as each plot represents the particular habitat. Each plot would be linked through both the decision of each farmer and the linkages between each farmer. The unit of analysis can be the "seed lot" as proposed by Louette, that seed which is selected by a farmer from an overall harvest for planting in the following year. Over multiple time periods time the seed lot selected each year would represent a crop population, and the model can link social and economic factors to this selection behavior. Finally the unit of analysis could be a variety, at a community or regional scale, known by a common set of characteristics or exchanged under a common name.

Migration parameters

The rate of migration and colonization in the case of a crop population would be the rate at which farmers acquire new seeds from within their seed network. Ideally there would be two main questions, how often does a farmer seek new seed of the same variety, and how often does a farmer seek new seed of a new variety. Conversely it will be important to model how long each farmer is able to depend on own saved seed, and how often is seed of the same variety periodically replaced. An obvious extension for the modeling approach is to extend the household-farm model to the seed replacement decision. The age of a seed lot is something that has been modeled analytically by CIMMYT in looking to extend Green Revolution gains (Brennan and Heisey), but there is a lack of empirical studies using household level data to model replacement seed demand. Furthermore there can be more complicated seed behavior, such as partial seed mixing to improve local materials or adapt improved foreign materials. In parallel to sequential adoption decisions, or multiple-variety in situ decisions, the model could introduce the same socio-economic variables such as market constraints or transactions costs to farmer seed decisions.

There are competing social and economic influences on the age of the seed that a farmer maintains. On the one hand an older seed lot may be grown by a farmer who has a larger plot, who has a wealth buffer that can allows survival of a low harvest without eating the saved seed, or skillful management of a population to maintain genetic integrity. On the other hand there are agronomic, genetic and seed health reasons would lead a farmer to periodically renew their seed from an external source. There is a need to explicitly model this aspect of farmer seed systems in order to understand the recycling of hybrids and the secondary diffusion impacts of modern varieties. This model will also be needed in targeting conservation efforts to ensure policies can target diversity-related crops and not modern varieties that are just managed as traditional varieties through informal channels.

Geneflow and Sampling Behavior

The flow of genes through pollen (and to a lesser extent uncontrolled seed dispersal and mixing) is another important area in which to model the interactions between human behavior and biological outcomes. The idea of pollen flow between landraces and modern variety has been a controversial topic within debates on field releases of GM crops; the biological science of geneflow under field conditions is not well documented. In fact, field trials of GM maize in Mexico were first suspended due to lack of scientific data on geneflow to inform the policy decision. Farmer behavior in traditional seed systems drives population biology parameters impacting geneflow: e.g will novel genes be selected for or be lost, can farmers perceive the geneflow and does their behavior take it into account?

These sorts of questions can only be answered through the characterization of farmer methods for selecting seed, and a methodological and empirical understanding of how farmer selection impacts local populations. The directional evolution of a crop population can start from the sample of the total population that is selected by the farmer for seed, and the characteristics farmers use to make selection decisions. These behavioral decisions will drive the parameters with fundamental conservation population outcomes like genetic drift, the degree of inbreeding, and the inclusion of deleterious mutations in a population. These parameters are also linked to the meta-population framework through the exchange networks – what quantity of seed is exchanged, and is each exchange a significant genetic bottleneck? Furthermore there are confounding effects, such as the loss of stored seed to pests, or the limited farmer understanding of sexual reproduction, that may impact biological outcomes but through indirect behavior.

Numerical and Empirical Examples

In the following section some of the parameters of a meta-population model will be explored through a combination of modeling and statistical analysis. The data are from an original household survey in a rural area of Puebla, Mexico, called the Sierra Norte de Puebla (SNP). The survey questionnaire had basic questions on seed management, household and market characteristics and thus provides the opportunity to begin to illustrate some of the aspects of the farmer based seed system. The crop modeled is maize, mostly a major white variety with some with color variants, and intercropped beans squash and other greens. The data and methods are more thoroughly described in VanDusen (2000).

In looking at the issue of geneflow, it may be useful to look at a stylized representation of pollen drift. In Figure 1 there are three fields, representing fields planted to three different varieties of an open pollinated crops such as maize.

[Figure 1]

The basic point here the difference between the impacts on Field 1 and on Field 3. The differences in the impacts of pollen flow are shown by comparing the ratio of border area to total area for each plot. This kind of geneflow can depend on many things, especially timing of pollen shed, and the shape and dimensions of the plots, which can be empirically measured to improve the model. (Louette).

The simplification in Figure 1 can be used to generate some numerical examples. A basic calculation combines typical farmer parameters of planting density, distance between border rows, and a biological assumption that the first two rows could be contaminated by a pollen from an adjacent field. This allows us to calculate a ratio of the two border rows as a percent of the total population contained in a field of a given size.

Border Area = Number of plants in two outer rows / Number of plants in total number of rows

This formula yields a simple nonlinear relationship between total area and a linear border. The obvious result is that smaller farmers have a much higher percent of their crop affected by geneflow.

Next we can combine this calculation for the impacts of geneflow with population level data for average plot sizes. Looking at household data from SNP in Figure 2 we can see that actually a large number of households have small enough plots to be affected.

[Figure 2]

In this figure the solid bars are a histogram of the percent of households with maize plots of a given size class. The line represents the calculation above on the percent of a maize plot of that size that could be contaminated by pollen from adjacent fields. For farmers with the principal maize plot up to 0.25 ha, at least 12% of the field could receive pollen, for farmers with up to half a hectare, 9% of the field is potentially contaminated. However in this sample, farmers with up to 0.25 ha represent 21% of the population and farmers with up to 0.5 ha represent (cumulatively) 55% of the population.

This calculation is not very precise or accurate, it can be improved by incorporating a statistical distribution on geneflow, by improving assumptions and data on field size and shape, etc. The point is to illustrate that the smaller the field, the more potential pollen geneflow from surrounding fields. This type of information could be improved by cross-referencing with GPS and GIS spatial layers in order to gather field level information on biological factors influencing geneflow.

Selection Behavior

Although it is more difficult to tie to the genetic outcomes, the seed selection behavior of farmers must be characterized in order to model the evolutionary pressures on farmer crop populations. The history of crop domestication, and the diversity of populations that farmers have created over time can be understood as the cumulative effect of every time a farmer has selected from one harvest to plant the next cycle. A few basic characteristics of sample behavior are presented in Table 1.

[Table 1]

First of all, in the SNP almost no farmers select in the field, which means that they cannot relate the seed that they choose to any of the characteristics of the plant that yielded the seed. Many of the efforts to improve the ability of farmers to breed their own seeds focus on this point, teaching techniques of field selection (As mentioned in the introduction , the recent renewed interest in Farmer Varietal Selection and Participatory Plant Breeding can be seen as seeking to incorporate or adapt different aspects of farmer seed systems). The second set of data show that a minority of SNP farmers are selecting their seed from the total population harvested. The majority, 55%, reported selecting seed at the time of planting. As the households are consuming maize during the months between harvest and planting, and many households produce at below their annual consumption level, the seed selection could be from only 50% of the harvested population. The third set of data in Table 1 are simply which members of the household make the decisions on which seeds to select. This can be interesting because of the traits are they selecting for. In some studies there is an idea that women select for consumption traits – culinary/food qualities, and men may select for production qualities. This type of information makes it possible to relate gender and cultural issues to crop populations.

The size of the selected seed lot is also a fundamental parameter in the viability of each sub-population of the overall population. In looking at maize populations, the effective population size is determined by the number of ears selected to make up a seed lot. Louette discusses a calculation made by Crossa that a minimum of 40 ears is necessary to capture 97% of the genetic variation of a maize population. A numerical calculation from the number of ears planted to an imputed area planted would require the planting density, the seed density, the ear size and density. This calculation could be used to represents the number of distinct individuals in a resulting planted area. The result is a linear relationship between area planted and number of ears used to select seed for that area. The results of this calculation are shown in Figure 3, and again are compared to the population of maize area planted in the SNP.

[Figure 3]

In this graph, the line corresponds to the number of ears that would correspond to a given sample size. The bars again represent the number of households whose primary maize plot corresponds to a given range of area. In the households from the SNP, the genetic threshold of 40 ears would affect 37.5 percent of the population. Below this threshold the consequence over repeated selection would be inbreeding and the resulting accumulation of deleterious mutations. The assumptions in generating the population size in this graph are limiting, but it is an example of how to calculate the households with dangerously low populations sizes.

In motivating the crop metapopulation it is also interesting to compare with Figure 2. In Figure 3 the smaller field has a population that may suffer from too low a genetic sample, while in Figure 2 the smaller field has population with increasing levels of genetic inflow. If these two effects are simultaneous, in Mexican maize populations the level of geneflow may compensate for the inbreeding problems of small populations (see discussion in Berthaud 2002).

Age and Origin of Seed

The design of a model of a crop metapopulation requires data on the spatial fragmentation of the landscape and the rates of migration and colonization. The primary household data from the SNP was gathered at the household level without the detailed information on flows between farmers for an explicit metapopulation. However we can explore the household data for information measuring these flows, and how to draw dynamic parameters from cross-sectional data. The primary source of information to look at from the SNP is the age of the seed that farmers are using. In Table 2 the age of the seed for maize varieties is presented, overall and separated by color type.

[Table 2]

This table alone can motivate population issues in the basic conservation questions for on farm conservation. Overall, and within the each variety, over half of the farmers reported having the seed for over 25 years. This indicates that the variety is locally adapted and farmers are not changing their traditional seeds for modern varieties. Within the population age structure for the white maize it appears to be relatively evenly balanced, that some farmers are changing seed in any given year, 19% in the last five years, 34% in the last ten years, and 43% in the last 15 years. The blue variety is much more scarce with only 19 out of 220 households growing it. The population age structure for the blue variety also appears more unbalanced, a bimodal distribution with 63% of farmers having seed for over 25 years and 21% replacing in the last five years. The yellow variety appears to have an intermediate position between white and blue.

In a crop metapopulation model it is also important to describe the source of seed, how far does seed move between the nodes or patches that farmers represent in the adaptation of the model. The survey data does not ask the specific source, but whether it was from the farmer's father, from a neighbor within the same village, or from outside of the village. These parameters could be used for an abstract model, or could provide the structure for how to incorporate explicit spatial effects. In Table 3 the seed source is presented for each color variety of maize.

[Table 3]

A similar pattern emerges to the age data above on age. The major white variety shows some balance between sourced from within the family (45%) and from other members of the community (52%). The blue variety is more unbalanced – 68% of farmers sourcing from their father. In looking towards the construction of the metapopulation model, the rate of exchange from outside the village, or between villages, appears to be around 4%. While this is a low number, cumulative over many years it could have a significant effect on the local populations.

To see if the age and sources correspond, the data on age of seeds is presented within cross tabulations with the source. Table 4 presents the cross tabulation for seed of the major variety, white maize.

[Table 4]

The age of the seed lot does appear to be correlated with the source of seed. The farmers who reported sourcing seed from their parents are more likely to have seed lots over 25 years old (71%). Looked at by age, those with seed for longer than 25 years are more likely to have seed from their parents (62%) than another member of the village or from outside. Conversely, the farmers with newer seed (acquired in the last five years) are more likely to have sourced seed from other villagers (86%) than from their family.

Finally it is necessary to look at what is perhaps a methodological question in determining the age of seeds, but has population questions. The sample survey asked two questions, first the age of the seed used by a farmer when it was being described generally. Later in the survey in a section on seed management, there was a question on "when was the last time that you had to replace or renew the seed". The answers were not consistent and Table 5 presents the results in a crosstabulation between the two questions.

[Table 5]

While at first only 43 farmers said they had acquired their seed in the last five years, later 119 stated that they replaced or renewed in last five years. Of the 114 farmers reporting seed over 25 years old, it appears that only 51, or less than half actually replanted their own seed for that long. The upper off-diagonal elements of this matrix, where farmers reporting renewing seed at a longer interval than the reported age of the seed, indicate problems with the data. On the other hand the lower off-diagonal elements could indicate that while farmers feel they are still planting the same variety, they can renew it with seed from another farmer but still be planting part of the same population. These apparent contradictions indicate problems with collecting this type of data, but can give us new information on how farmers understand the populations they are managing and indicate ways to model seed replacement separately from varietal replacement.

Regression for Age of Seed

The motivation for modeling the farmer seed system as a crop metapopulation is to derive policy relevant information. The next step for the model is to explore how to link the key parameters of the crop metapopulation to the socio-economic context. Individual farmers are modeled as economic agents who act according to social conditions, household demographics, local constraints and endowments, and market conditions. Model specification is not clear from theory or empirically, but there are a range of different forces impacting a household's participation and use of a local seed system. The approach used is to nest a series of competing models to look at the performance of each and then a joint model (Smale, Just and Leathers)

The meta-population model gives two obvious starting points for the econometric analysis: the presence or absence in a given patch, and the movement between patches. The first question, whether a farmer grows a specific variety is analogous to the household model used in VanDusen (2000), which looks impact of socio-economic variables on the probability that a household grows a diverse set of landraces of maize, and of other milpa crops. The second question, how often a farmer replaces a seed, or the age of a farmer's seed, entails an econometric exploration of the factors driving the differences in ages of seed described in the section above. An econometric analysis of the age of a farmer's seed is the focus of the empirical section of this paper.

The underlying household-farm model is the same: a random utility model of a farmer's decision to replace a seed lot. Economic theory predicts that, given perfect markets, the decision to replace seed will be solely due to the price and yield of the output. As mentioned above there are agronomic conditions for which the quality of a seed lot may decline. The price is expected to be the same for all households, and only vary by village specific transactions costs, which mainly drive differences in output markets and key inputs such as labor. The isolated and remote developing country contexts of the SNP create conditions where markets may not work well. Imperfect markets introduce to the model variables on distance to market and labor market conditions when households use endogenous prices for both factors and outputs.

As mentioned above the independent variable has already been presented in Table 2. The summary statistics for the explanatory variables are presented in Table 6.

[Table 6]

The variables are grouped into three sub-groups, household characteristics, farm characteristics, and market conditions. Household characteristics include age, family size, and two variables that serve as indicators of indigenous conditions, one a village level variable constructed from the percent of villagers speaking an indigenous language and a second based on a traditional planting practice in the region. The agro-ecological variables include the area planted, location in high altitude zone, the total number of species grown in a maize plot (milpa), and the total number of plots planted to milpa. The market conditions are a village level variable on the use of hired labor in maize production, the distance from the nearest regional market, and whether the household has a member who is a national or international migrant. A final variable is a proxy for household wealth, which was composed by ranking the size and material of a household's dwelling.

The specification used for this model is a Tobit, because the data are right censored at 30 (farmers with seed older than 30 years are grouped together, the variable is continuous otherwise). The model was run for each subset of household, farm, and market conditions and then jointly for all three. The results are presented in Table 7

[Table 7]

In the regressions on just household characteristics only age is significant– being more indigenous does not appear to impact the age of seed that a farmer holds. In the regression on farm characteristics, the total area, being in the higher altitude area, and the total number of varieties planted all increase the probable age of a farmers seed. The area impact was discussed above, a larger area reducing the probability of inbreeding problems. The high altitude area is more dedicated to maize production; the lower altitude area has more alternative activities such as coffee. That the total diversity of the milpa is positively correlated to the number of species has more of a conservation implication than explanation. A positive correlation would mean that farmers with older landraces and more diversity could be jointly targeted. In the regression on market conditions, only the coefficient on intra-Mexico migration is significant, but it is interesting that it has an opposite signs to the coefficient on US migration. Having a national migrant decreases the probable age of the seed, while having a US migrant might increase the probable age. This could be due to the fact that in the cases with intra-Mexico migration, the head of the household migrates and may lose seed when missing an agricultural cycle, while in US migration it is more common that a son is out of the household and subsidizing that a parent continues farming. Finally the coefficient on household wealth is positive and significant, a wealthier household is more likely to have older seed. A first explanation would be that a wealthier household has a greater ability to withstand weather or consumption shocks and is less likely to be forced to eat the seed lot saved for planting in the next season. In the joint model results only the coefficients for age and migration are significant. By theory these migration effects should be secondary to agronomic or technical conditions, and thus undermine the credibility of the results.

These econometric results are interesting but not particularly compelling that they are explaining the dynamics of seed loss and replacement. Several other specifications were fitted, starting from a logit model identifying specifically the households who reported having the same seed forever. Due to the fact that loss of seed over time can be modeled as the duration of a farmer's seed, semi-parametric and non-parametric specifications from duration models were tried. The Cox proportional Hazards model, appropriate for survival analysis, and the Han and Hausman ordered Logit model, with the dependent variable coded into general categories were also fit to the data. Unfortunately neither model was able to improve the explanatory power over the Tobit. One specification problem was that the approach was to use a common set of explanatory variables for the presence/absence model as the migration/colonization model, and this may have to be amended for more specific sets for each model.

In order to view these econometric results in a broader context, the results for age of seed lot are compared to regressions on the number of maize varieties planted and the total number of species planted in the milpa, including intercropped beans and squash. In previous work on in situ conservation and technology adoption, this has been a more common approach, to look for social and economic factors to explain the planting decisions of household farms (Meng, Bellon and Risopolous, Smale, Bellon and AguirreGomez). For comparison the regression was run in a single Tobit for each explanatory variable, with the same sets of explanatory variables. However total milpa varieties is removed from the explanatory variables and made a dependant variable, which changes the results for seed age slightly. The results are presented in Table 8.

[Table 8]

In the diversity equations explaining how many varieties or species planted, it is obvious that different factors are driving the outcome. For maize, the positive and significant signs on living in an indigenous village means that this increases the probability that a household plants a minor, colored variety. For both maize and total milpa species, where the household carries out a traditional planting practice increases the probability of planting a more diverse milpa. One of the more interesting policy results is the positive and significant sign on the coefficient for the distance to market. The household situated in village further from the regional market center is more likely to maintain a more diverse milpa, but it is interesting to note that this important market variable was not significant in explaining seed age. Finally the national migration to Mexican destinations appears to decrease the probable number of maize varieties sown, just as it decreased the probable age of seed lots in the previous section. The first conclusion from these comparisons is that some of the forces affecting planting decisions appear to be different than those affecting seed age. This offers challenges to the modeling, that the different dimensions of diversity management operate with different underlying processes, and that a common set of socio-economic characteristics may not be appropriate.

Conclusions

The outcome from this paper is not a set of policy recommendations, but a proposal of a methodology and suggestions on how to implement it. The benefits are a more robust understanding of the factors influencing CGR conservation, and a means to quantify and measure the impacts of farmer seed systems. If this approach can be fully developed, both in terms of gathering the appropriate data and constructing a simulation component, the outputs could be used for several urgent policy questions. This paper has been basically focused on the conservation of traditional varieties of crops, and how we can extend information about the households to incorporate information on variety management that may be crucial to conservation decisions. However the approach has been to model it in a manner that is flexible enough to address other types of policy questions that need information on the impacts of farmer seed systems.

First is to model how an individual farmer's behavior will impact the local and regional impact of biosafety decisions to release GM crops. If there is GM contamination of local farmer varieties, understanding rates of exchange, extent and efficiency of exchange networks, and interactions with other crop decisions are fundamental. The second is to model potential impacts of participatory breeding and selection schemes that propose to utilize farmer exchange to diffuse varieties. Many novel approaches seek to use farmer seed systems, but a quantification of these systems, a comparison to formal or commercial networks, and the robustness under disaster or crisis situations is necessary. Third, the output is to improve understanding of how farmers use informal methods to access improved CGR. Where gains from crop breeding can be accelerated, understanding farmer selection behavior, varietal turnover, and interaction with economic forces can be used to target breeding strategies.

Directions for Future Research

This is only a beginning for how to adapt meta-population models to farmer seed systems. The immediate next step is to improve the analysis of seed loss and turnover, both analytically and through improved regression analysis. Explaining the decisions to change varieties, or change the seed of a given variety through household level data would fill a gap in the existing literature. The intermediate step is to find appropriate data sets where the questions on seed management and seed exchange can be matched to socio-economic data about the households. A series of recent studies have gathered this type of data and are in the process of preliminary analysis. For the future the goal will be to design new household surveys in areas of CGR interest to quantify and analyze farmer seed systems.

The next methodological step is to take parameters derived from household level data, and to construct a simulation model to get at inter-temporal impacts of different conservation scenarios. The application of meta-population models to population viability analysis is used in ecology to make decisions on the allocation of resources in conservation of habitat. The application to crop metapopulation would not be limited to how to conserve, but would similarly strive to provide the ability to forecast population and genetic outcomes for changes in the environment, in this case the economic and social variables affecting farmer behavior. The empirical basis for the simulation model will suffer from brutal and painful simplification in design and construction, but for the important parameters will rely on results from econometric models that can be recovered and adapted from different case studies. Finally the model will hopefully be able to incorporate some stochastic components, to be able to add the dimension of risk to policy information based on costs and benefits.

References:

Aguirre Gómez, J. A., M. R. Bellon, and M. Smale. "A regional analysis of maize biological diversity in Southeastern Guanajuato, Mexico." *Economic Botany* (2000): v, 17.

Almekinders, C., D. Louette "Examples of innovations in local seed systems in Mesoamerica" in Almekinders C, and DeBoef W." *Encouraging Diversity : The Conservation and Development of Plant Genetic Resources* PG. 219-222. 2000 [Figures]. .

Ashby, J. A., and L. Sperling. "Institutionalizing Participatory, Client-Driven Research and Technology Development in Agriculture." *Development and Change* 26, no. 4(1995): 753-770.

Bellon, M. R. "The Dynamics of Crop Infraspecific Diversity - a Conceptual Framework At the Farmer Level." *Economic Botany* 50, no. 1(1996): 26-39.

Bellon, M. R., and J. Risopoulos. "Small-scale farmers expand the benefits of improved maize germplasm: A case study from Chiapas, Mexico." *World Development* 29, no. 5(2001): 799-811.

Bellon, M. R., and J. E. Taylor. ""Folk" Soil Taxonomy and the Partial Adoption of New Seed Varieties." *Economic Development and Cultural Change* 41, no. 4(1993): 763-86.

Berthaud, J., et al. (2002) Farmers management of maize landrace diversity. A case study in Oaxaca and beyond. Beijing.

Brush, S. B. (1999) Genetic Erosion in Centers of Diversity: A revision, ed. FAO-WIEWS. Prague.

Brush, S. B. "In Situ Conservation of Landraces in Centers of Crop Diversity." *Crop Science* 35, no. 2(1995): 346-354.

Byerlee, D., and P. W. Heisey (1990) Wheat Varietal Diversification over Time and Space as Factors in Yield Gains and Rust Resistance in the Punjab, ed. P. W. Heisey, vol. CIMMYT Research Report No 1. Mexico, DF, CIMMYT.

GTZ (2002) "Options for Supporting On-farm Conservation in Eastern and Southern Africa".

Hanski, I. "A practical model of metapopulation dynamics." *Journal of Animal Ecology* 63, no. 1(1994).

Heisey, P. W., and J. P. Brennan. "An analytical model of Farmers' Demand for Replacement Seed." *American Journal of Agricultural Economics* (1991).

Heisey, P. W., et al. "Wheat Rusts and the Cost of Genetic Diversity in the Punjab of Pakistan." *American Journal of Agricultural Economics* 79, no. 3(1997): 726-737.

Heisey, P. W., et al. "Varietal Change in Post-Green Revolution Agriculture - Empirical Evidence for Wheat in Pakistan." *Journal of Agricultural Economics* 44, no. 3(1993): 428-442.

IPGRI (2002) American Seed Project - Farmers seed system, vol. 2003.

Jarvis, D. I. *A Training Guidede for In situ Conservation On-Farm*. Rome: IPGRI, 2000.

Joshi, A., and J. R. Witcombe. "Farmer participatory crop improvement .2. Participatory varietal selection, a case study in India." *Experimental Agriculture* 32, no. 4(1996): 461-477.

Kaplinsky, N. "Conflicts around a study of Mexican crops - Reply." *Nature* 417, no. 6892(2002): 898.

Kaplinsky, N., et al. "Biodiversity (communications arising): Maize transgene results in Mexico are artefacts." *Nature* 416, no. 6881(2002): 601.

Louette, D., A. Charrier, and J. Berthaud. "In Situ Conservation of Maize in Mexico - Genetic Diversity and Maize Seed Management in a Traditional Community." *Economic Botany* 51, no. 1(1997): 20-38.

Louette, D. "Traditional management of seed and genetic diversity: what is a landrace?" in Brush SB." *Genes in the Field : On-Farm Conservation of Crop Diversity* Pg. 109-142. 2000 .

Louette, D., and M. Smale. "Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico." *Euphytica* 113, no. 1(2000): 25-41.

McArthur, D. R., and E. O. Wilson. *The theory of Island Biogeography*. Princeton, NJ: Princeton Univ Press, 1967.

Meng, E. C.-H. "Land allocation decisions and In situ conservation of crop genetic resources : the case of wheat landraces in Turkey." 1997.

Perales Rivera, H. R. "Conservation and evolution of maize in Amecameca and Cuautla valleys of Mexico." 1996.

Quist, D., and I. H. Chapela. "Maize transgene results in Mexico are artefacts - Reply." *Nature* 416, no. 6881(2002): 602.

Quist, D., and I. H. Chapela. "Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico." *Nature* 414, no. 6863(2001): 541-543.

SADC-FANR (2001) Seed Security Network, vol. 2003.

Smale, M., M. R. Bellon, and J. A. A. Gomez. "Maize diversity, variety attributes, and farmers' choices in Southeastern Guanajuato, Nexico." *Economic Development and Cultural Change* 50, no. 1(2001): 201-225.

Smale, M., R. E. Just, and H. D. Leathers. "Land Allocation in Hyv Adoption Models - an Investigation of Alternative Explanations." *American Journal of Agricultural Economics* 76, no. 3(1994): 535-546.

Sperling, L., and M. E. Loevinsohn. "The Dynamics of Adoption - Distribution and Mortality of Bean Varieties among Small Farmers in Rwanda." *Agricultural Systems* 41, no. 4(1993): 441-453.

Sthapit, B. R., K. D. Joshi, and J. R. Witcombe. "Farmer participatory crop improvement .3. Participatory plant breeding, a case study for rice in Nepal." *Experimental Agriculture* 32, no. 4(1996): 479-496.

Tripp, R. "Can biotechnology reach the poor? The adequacy of information and seed delivery." *Food Policy* 26, no. 3(2001): 249-264.

Tripp, R. *New seed and old laws : regulatory reform and the diversification of national seed systems*. London: Intermediate Technology Publications on behalf of the Overseas Development Institute, 1997.

Witcombe, J. R., et al. "Farmer participatory crop improvement. IV. The spread and impact of a rice variety identified by participatory varietal selection." *Experimental Agriculture* 35, no. 4(1999): 471-487.

Wright, M., and M. Turner (1990) Seed Management Systems and Effects on Diversity, ed. D. Wood, and J. M. Lenne. Wallingford, CABI.

Figure 1 - Stylized Model of relationship between geneflow and planted area

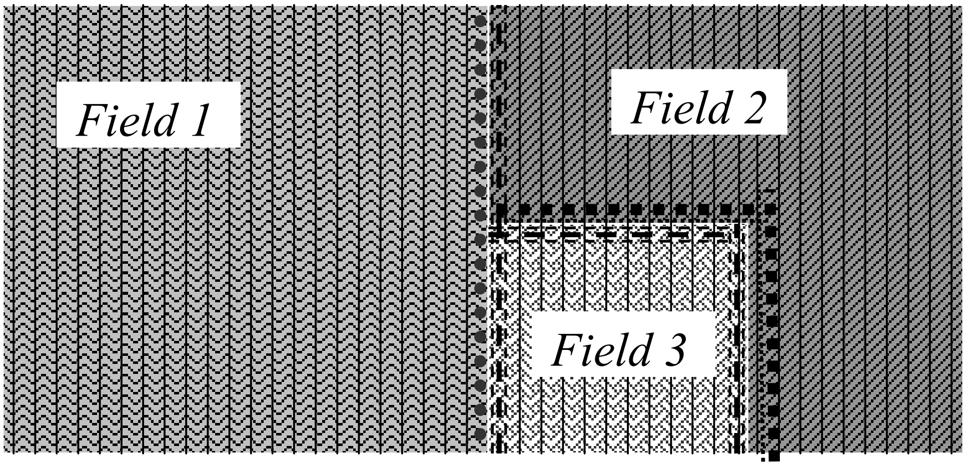

Figure 2 - **Numerical simulation of rate of geneflow contamination as a function of maize area plotted against percent of households by category of maize area**

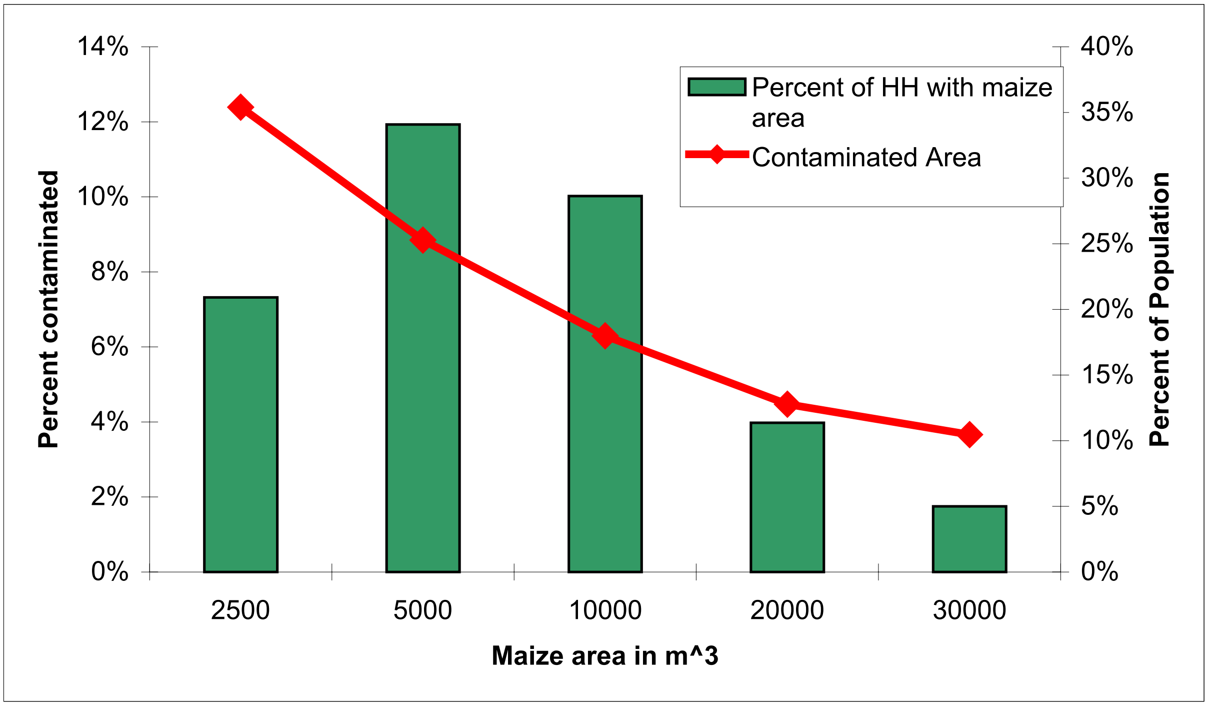

Figure 3: Number of ears selected to plant a given maize area plotted against the percent of total households with maize area

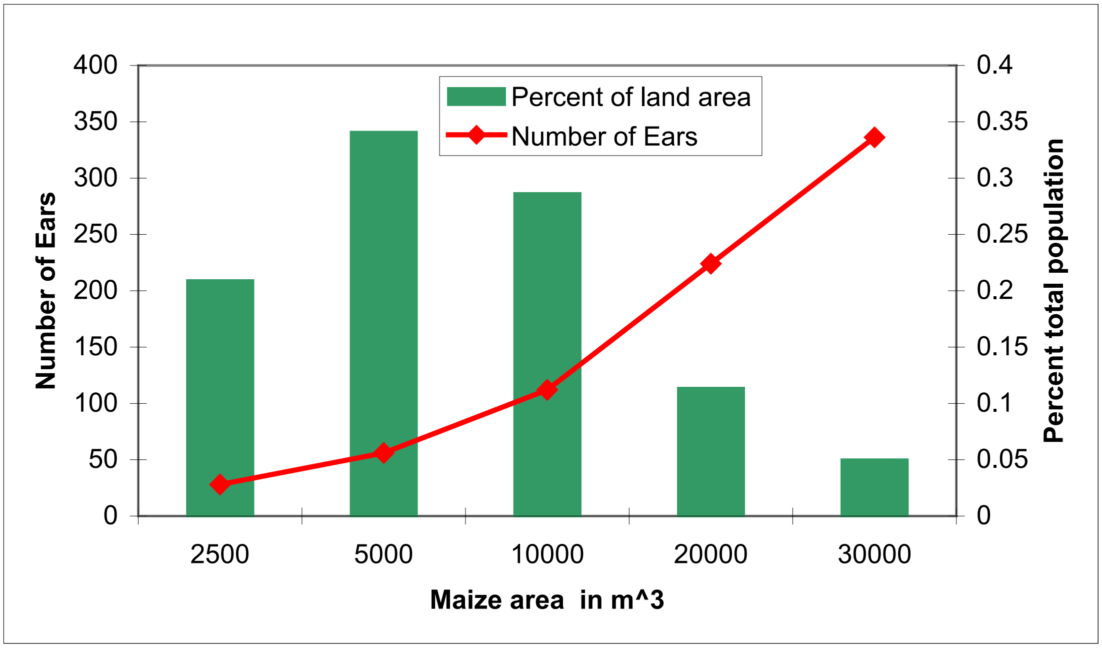

Table 1 – Selection Behavior in SNP survey sample

	Percent of Farmers
Select in field	% 1
Select in storage	% 99
Select at harvest	% 27
Select during the year	% 18
Select at planting	% 55
Selected by man	% 47
Selected by woman	% 20
Selected by both	% 33

Table 2 : Age of Maize Seeds, overall and by color type

Years	*White*		*Yellow*		*Blue*		*Total*	
0-5	*42*	*19%*	*11*	*23%*	*4*	*21%*	*57*	*20%*
5-10	*32*	*15%*	*2*	*4%*	*1*	*5%*	*35*	*12%*
10-15	*17*	*8%*	*4*	*9%*	*1*	*5%*	*22*	*8%*
15-20	*12*	*5%*	*2*	*4%*	*1*	*5%*	*15*	*5%*
20-25	*2*	*1%*	*1*	*2%*	*0*	*0%*	*3*	*1%*
>25	*115*	*52%*	*27*	*57%*	*12*	*63%*	*154*	*54%*
Totals	*220*		*47*		*19*		*286*	

Table 3: Sources of Seed by Color

	White	*Yellow*	*Blue*
Father	*45%*	*56%*	*68%*
Village	*52%*	*40%*	*32%*
Other	*3%*	*4%*	*0%*

Table 4: Crosstabulation of Age and Source of Seed Lot (White Maize)

	Father	Neighbor	Other	Totals
Age 0 to5	4	37	2	43
Age 5 to 25	24	36	3	63
Age > 25	71	42	1	114
Totals	99	115	6	220

Table 5: Crosstabulation of stated Age of seed lot and years since the last time that the seed was "renewed"

	Last time seed renewed			
	0 to 5	5 to 25	> 25	Totals
Age 0 to5	33	1	9	43
Age 5 to 25	27	23	13	63
Age > 25	59	4	51	114
Totals	119	28	73	220

Table 6 : Summary Statistics for Household level Variables

	Mean	Minimum	Maximum
Age HH Head	51.409	25	88
Family Size	5.250	1	10
Indigenous Village	0.412	0	1
Traditional Practice	0.273	0	1
Maize Area	0.954	0	7
Higher Elevation	0.459	0	1
Number of Plots	1.245	0	4
Total Varieties	2.991	1	9
Hired Labor	0.478	0.09	0.79
Distance to Market	7.259	0	15
US Migrants	0.077	0	1
Mexico Migrants	0.564	0	1
Wealth Index	6.538	0	20

Table 7: Set of Tobit Regressions on Age of Seed ,

	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	-0.445	-0.06	15.181	3.64 **	17.504	2.68 **	-17.637	-1.68 *
Age HH Head	0.533	4.43 **					0.510	4.28 **
Family Size	0.763	1.10					0.281	0.40
Indigenous Village	-7.211	-1.38					-1.978	-0.32
Traditional Practice	0.303	0.09					0.621	0.18
Maize Area			3.219	1.73 *			2.263	1.25
Higher Elevation			6.425	1.84 *			3.666	0.99
Number of Plots			0.869	0.47			0.222	0.12
Total Varieties			1.931	1.79 *			1.649	1.56
Hired Labor					7.969	0.68	9.530	0.84
Distance to Market					0.603	1.43	0.419	0.98
US Migrants					10.442	1.60	11.806	1.89 *
Mexico Migrants					-7.025	-2.20 **	-5.079	-1.67 *
Wealth Index					0.863	1.86 *	0.620	1.38
Sigma	19.880	12.67 **	20.408	12.65 **	20.542	12.65 **	18.789	12.75 **
Log-L	-546.313		-551.02		-551.2		-535.9	

Table 8: Set of Tobit Regressions - Number of Varieties planted, Age of Seed

	Total Maize		Total Milpa		Age of Seeds	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	-2.096	-1.84	0.227	0.27	-16.397	-1.57
Age of HH Head	-0.007	-0.52	0.008	0.88	0.523	4.35 **
Family Size	-0.064	-0.90	-0.007	-0.13	0.284	0.40
Indigenous Village	2.152	3.36 **	0.555	1.15	-1.338	-0.22
Traditional Practice	1.033	2.94 **	0.746	2.78 **	1.542	0.45
Area planted to Maize	0.545	3.25 **	0.265	1.90	2.677	1.47
Higher Elevation Area	0.602	1.70	1.245	4.40 **	5.203	1.44
Number of Plots	0.202	1.08	0.215	1.49	0.483	0.27
Hired Labor	-1.534	-1.33	-0.028	-0.03	9.724	0.86
Distance to Market	0.139	2.82 **	0.099	2.87 **	0.553	1.30
US Migrants	0.053	0.09	-0.477	-1.09	11.123	1.77
Mexico Migrants	-0.722	-2.28 *	0.099	0.42	-4.972	-1.63
Wealth Index	0.062	1.43	-0.003	-0.09	0.619	1.38
Sigma	1.555	8.48 **	1.647	18.53 **	18.940	12.74 **
Log-Liklihood	-154.735		-387.03		-537.130	