MATHEMATICA
Policy Research, Inc.

MEMORANDUM

P.O. Box 2393
Princeton, NJ 08543-2393
Telephone (609) 799-3535
Fax (609) 799-0005
www.mathematica-mpr.com

TO: Rebecca Tunstall

FROM: Ken Fortson; Anu Rangarajan

DATE: 3/4/2008
MCC-008

SUBJECT: Water-to-Market Evaluation Plan

The Millennium Challenge Account with Armenia (MCA) aims to increase household income and reduce poverty in rural Armenia through improved performance of the country's agricultural sector. Armenia plans to achieve this goal through an integrated, nationwide initiative to improve major components of the rural infrastructure, with a focus on lifeline roads and the irrigation system, including improvements in water management. By improving living standards among rural residents, these investments can in turn lead to future economic growth in rural areas and throughout the country as a whole.

To support the infrastructure rehabilitation efforts, MCA has contracted with ACDI/VOCA and their partners, VISTAA and Euroconsult—hereafter referred to collectively as ACDI—to implement Water-to-Market (WtM) activities that include training farmers in water management and high value agriculture, as well as credit and post-production services. A total of 60,000 farmers will be trained in water management practices over a five-year period, and of those, 30,000 will also be trained in high value agriculture.

MCC has commissioned a rigorous impact evaluation to separately examine each of the three main components of the MCA-Armenia program. This memo describes our design for evaluating the WtM activities. We will prepare separate designs for the road rehabilitation and irrigation infrastructure activities after the implementation plans for those activities have been finalized.

To maximize the rigor of the analysis, MCC instituted a random assignment design, whereby villages would be randomly assigned into a treatment group, whose farmers would be offered training, and a control group, whose farmers would not be offered training during the follow-up period. The evaluation will thus focus on assessing the impacts of the training components.

The key research questions guiding our design of the evaluation for the WtM training are:

- Did the program affect the irrigation and agricultural practices of Armenian farmers?

- Did the program affect agricultural productivity?

- Did the program improve household well-being for the targeted farmers, including income and poverty?

We first summarize ACDI's farmer training programs, followed by an overview of the random assignment design. We then discuss the main source of data, the Farming Practices Survey, followed by the Irrigation PIU data that could possibly be used as a supplemental data source. Lastly, we discuss in detail our econometric approach for estimating program impacts, and alternative specifications that can be employed to further explore the key research questions.

I. WATER-TO-MARKET TRAINING

ACDI will train 60,000 farmers over a five-year period. Participating farmers will be trained in a variety of techniques designed to help them improve the quality and quantity of crops, conserve water, and introduce new varieties of crops. The training will have two components, water management and high value agriculture (HVA). Additionally, farmers who participate in training will be permitted to apply for the agricultural credit programs. The post-harvest processing components will be implemented across the country, and thus, cannot be evaluated on their own. Therefore, the evaluation is designed to estimate the impact of the *marginal* training programs, that is, the impact of the training programs and access to credit on top of any other services and programs that are available to farmers in these villages, including any information disseminated widely through mass media.

Water Management. Sixty-thousand farmers will be trained in region-specific water management techniques. Because the large majority of Armenian farmers have small plots of land, averaging less than two total hectares, the focus of these training sessions will be to teach farmers low-cost irrigation technologies that are especially suitable for small-scale farming operations. These methods will help farmers to use water more efficiently, which can promote both cost savings (through water conservation) and increased quantity and quality of crops.

High Value Agriculture. Approximately half of the farmers who are trained in water management technologies will also be trained in higher-value agricultural methods. It will be implemented in the subset of villages with conditions that will enable them to utilize this training. An important objective of the HVA training will be increasing quantity and improving quality of the crops, as it is for the water management training, but a special emphasis will be placed on cultivating new, higher-revenue crops as well as higher-valued varieties of common crops, such as some organic versions. Training in new and improved crop varieties will be supported by demand for these crops, as evidenced by ACDI market studies and other marketing activities.

The specific training techniques for both components will be tailored to the agricultural conditions of the region, with appreciable variations across the four agricultural zones. Nonetheless, while the specific techniques will vary, the training practices are expected to be reasonably similar. ACDI anticipates that each training session will involve 20 to 25 farmers. Two to three days of the training will be theoretical lessons taught in a classroom setting; these classroom sessions will be supplemented with practical lessons taught on a demonstration farm that is set up and maintained for this purpose.

The demonstration farm is a critical part of the initial training and also can help reinforce the lessons even after training is completed. Farms are selected to serve as demonstration sites based on their proximity to other farmers in the village and the farmer's willingness to adopt new technologies and facilitate other farmers' understanding. ACDI provides equipment for the demonstration farms; in exchange, the demonstration farms are used as the site of the practical training, and also a resource where farmers in the village can go to see the technologies in practice, beyond the official training session.

Demonstration farms will serve anywhere from one to five villages, depending on the number of eligible farmers in those villages and their proximity. Some demonstration farms exist from previous activities, and these will be supplemented with new demonstration farms. Most demonstration farms will be new. The trainers will be agricultural experts who come from the same region so as to ensure that they are familiar with the local climatic and agricultural conditions, and so that they are available for technical assistance beyond the training sessions. Once a demonstration farm is established, ACDI will provide several rounds of training at that demonstration farm to saturate the associated villages as much as possible, because high participation rates will maximize the use of a single demonstration farm.

Access to Credit. Farmers who complete training will be permitted to apply for agricultural credit, increasing their financial means to invest in equipment and supplies to implement the techniques taught in the training. They will still need to apply and be approved for credit. This feature of the program is particularly important given that many of the techniques advocated in the training require expensive equipment that might not otherwise be affordable.

Outreach and Enrollment. For both types of training, coordinators will target farmers who are members of Water User Associations (WUAs). They will use posters to advertise at village centers and work with mayors to mobilize farmers to participate. In some villages, mayors may also be able to identify farmers who are particularly likely to participate, and these farmers could be targeted for additional recruitment efforts.

Rollout of Training. ACDI plans to train 2,000 farmers in the pilot phase (2007), 13,000 in the second Compact year (Oct. 2007 through Sept. 2008), 19,000 in each of the third and fourth Compact years (Oct. 2008 through Sept. 2010), and 7,000 farmers in the final year of the Compact (Oct. 2010 through 2011). By the second year, ACDI will have trained at least some farmers in all of the four agricultural zones. Training in the earlier years can only be implemented in villages that have adequate water sources to begin with. Many villages will have Compact-funded irrigation system rehabilitation efforts that will greatly improve the water system, at which point training will be more fruitful.

II. RANDOM ASSIGNMENT

The ideal method for separating program influences from other factors is to compare outcomes for the group who are provided the intervention with the outcomes for the same group if they were not offered the intervention. However, once persons or communities are offered the intervention, it is not possible to know what their outcomes would have been if they were not given the opportunity to participate. It can only be approximated by comparing their outcomes to the outcomes of some other group.

Random assignment, the most rigorous way to measure program impacts, is frequently referred to as the gold standard of evaluation designs. Essentially, when implemented carefully, random assignment leads to the creation of two virtually identical groups at baseline, with the only difference being that only one group (the treatment group) is exposed to the intervention, while the other group (the control group) is not. As a result, any changes observed between the two group over time can be attributed to the effects of the intervention with a known degree of statistical precision.

Unit of Random Assignment. Ideally, we would randomly assign individual farmers to receive training or not, and compare outcomes for the two groups. However, because these training sessions are community-level interventions, making it difficult to exclude individual farmers, such an approach is not feasible in this context. Our basic approach is to randomly assign villages to the treatment group of farmers, who are eligible for on-farm water management training and HVA training, or the control group of farmers, who are not. ACDI has grouped villages into clusters. Most clusters include only one village, but some include as many as five villages that are in close geographic proximity. All farmers who are WUA members and live in a cluster of villages selected for the treatment group will be permitted to participate in water management training. Farmers who are in the control group of village clusters would not be offered water management training until several years later. All villages will ultimately be provided training, and random assignment is used to determine *when* they are offered training.

Randomly assigning entire village clusters in this way, rather than individual farmers or villages, guards against contamination of the control group—the possibility that control group members get the same services as the treatment group. There are two types of contamination. The first type of contamination is if farmers in control group villages nonetheless participate in training. This could be problematic if control group members hear about the training activities and show up to training themselves. A different type of contamination could occur if farmers who participate in training teach farmers in the control group about the techniques they learned. Either of these types of contamination would be problematic for the evaluation because we would be unable to compare those who were offered training to those who were not offered training; with contamination, both the treatment and control group have access to or benefit from training. Generally, ACDI has chosen village clusters that are sufficiently far apart geographically to ensure that there is little chance that farmers in a control group village cluster would either participate in the training or learn about the water management techniques through other means.

However, in some areas—particularly the Ararat Valley region—many villages are located in close proximity. While we cannot completely eliminate the possibility of contamination here, it will be important for the planned implementation to strive to avoid such contamination problems by, for example, ensuring that recruiting techniques for the training attract treatment group farmers without influencing control group farmers. The WtM training program exit questionnaire will also inform us about where farmers reside, which will help us assess the extent to which control group farmers are "crossing over" and receiving training in spite of being randomly assigned to not be eligible.

Implementing Random Assignment. Random assignment was conducted for the subset of villages that have adequate water and could potentially be served early in the Compact. We randomly assigned villages to one of three groups: those who would be served in the second year of the Compact; those who would be served in either year 3 or year 4 of the Compact; and those who can be served in the final year of the Compact. The earliest group constitutes our treatment group, and the latest group our control group—impacts will be measured after the treatment group has been provided training but before the control group has. The middle group, those who are served in the third or fourth year, will not be included in the impact evaluation. Only villages that were considered ready for WtM training were included in the randomization; some villages currently have poor sources of water, and thus, would not benefit from training until their irrigation systems are rehabilitated. Such villages may receive training in the future, but they will not be included in the impact evaluation. We also excluded from the random assignment all villages that were included in the pilot phase of the WtM training or where ACDI has already developed demonstration farms.

Random assignment was conducted within strata defined by WUAs to preserve regional balance, which created balanced treatment and control groups along this dimension. The distribution of villages by treatment status for each agricultural zone is reported in Table 1. The probability of a village being assigned to the treatment group was approximately the same for all WUAs, with most of the deviations occurring due to rounding. An exception, however, is the Mountainous Zone, where a smaller proportion of villages were selected for the research groups (years 2 and 5), while most villages were assigned to the nonresearch group. This zone was undersampled largely because MCC anticipates very low impacts, so the evaluation will focus more on the other zones where MCC is more optimistic about the prospects for improvement. A total of 120 clusters were assigned to the treatment group and 80 to the control group, with these 200 clusters containing 223 villages in total. (For simplicity of exposition, we hereafter refer to village clusters as "villages.")

Table 1. Distribution of Village Clusters by Year of Training and Agricultural Zone

	Ararat Valley	*Pre-Mountainous*	*Mountainous*	*Sub-Tropical*	***Yearly Total***
Year 2: Treatment	44	58	12	6	**120**
Years 3 and 4: Nonresearch	18	19	38	2	**77**
Year 5: Control	28	38	10	4	**80**
Total	90	115	60	12	**277**

III. FARMING PRACTICES SURVEY

The Farming Practices Survey (FPS) will serve as the primary data source for the impact evaluation. Approximately 25 interviews will be completed in each of the villages in the treatment and control groups, with fewer in the smallest villages and more in the largest. MCA-Armenia has contracted with AREG to field the FPS in the first, baseline year, and the FPS will subsequently be conducted each year of the follow-up period, at the end of 2008, 2009, and 2010.

Sample Frame. With the help of village mayors, the FPS targets the households of farmers who are most likely to benefit from the training programs: those who are actively engaged in farming and have been tied to the community for several years. These farmers are identified through an iterative process. First, MCA-Armenia requested that the WUAs work with village mayors to compile a list of farmers meeting our specific criteria in each village. The number of farmers requested depended on the size of the village, but averaged about 60. The sample was then drawn from these lists.

Pretesting revealed that these lists were of mixed quality, however, often because the WUAs had not consulted with the mayors in compiling them. Thus, the sample was updated with the assistance of village mayors and marz officials, either at the marz offices or in the village itself. The mayors reviewed the lists to determine whether the farmers indeed met our criteria. If an insufficient number of farmers from the lists were eligible—that is, in cases where the WUA had failed to consult with the mayor—then the mayor helped AREG update the list in accordance with our survey eligibility criteria.

Intermediate Outcomes. While most of the outcomes of primary interest to MCA-Armenia and MCC are longer-term outcomes, such as economic improvements, these outcomes may not be immediately observable. Consequently, we will closely examine intermediate outcomes through which the training programs are intended to improve household income. We would expect an impact on households' income only if we observe that a substantial proportion of the targeted farmers are actually participating in training, and perhaps most importantly, are then applying the techniques they learn. Examining the intermediate outcomes also establishes the counterfactual—what services the villages would have received and what practices they would have adopted even in the absence of the WtM programs. Table 2 summarizes the key intermediate that can be examined using the FPS data.

Table 2. Intermediate Outcome Measures

Intermediate Outcome Measures	Time Frame
Participation in Agricultural Training. Whether attended any irrigation or agriculture training (including training sponsored by other sources); type of training attended (e.g., classroom, video, or practical); whether received a certificate indicating the full training was attended.	Last Year
Adoption of HVA and Irrigation Practices. Which irrigation practices were used, focusing on those taught in training sessions; whether those practices had perceived time or labor savings.	Last Agricultural Season
Investment in Agricultural Technology or Equipment. Ownership of personal reservoir or water pump; ownership or rental of trucks, tractors, combines, seed planters, and sprayers.	Last Agricultural Season
Cropping Patterns. Specific crops grown, especially high-value crops; amount of land devoted to cultivation of each crop; total hectares of land devoted to crops; whether household cultivates a kitchen plot; reason(s) for changes in cropping patterns.	Last Agricultural Season

Final Outcomes. The ultimate goal of the MCA-Armenia programs is to increase household income in rural Armenia, and hence, these outcomes are an important focus of the FPS instrument. Because a full accounting of all sources of household income would require far longer to administer than the allotted time for each interview, the survey concentrates on sources of income that are most directly affected by the training programs, specifically, income from agricultural production and from employment by the farmer and his or her immediate family. We can also use the average sale price of specific crops for other farmers in the village to monetize crops that are consumed by the household or bartered. Additionally, the FPS asks for estimates of expenditures on key categories of consumption, and for income from other sources. Table 3 summarizes the key final outcomes that can be examined using the FPS data.

Table 3. Final Outcome Measures

Final Outcome Measures	Time Frame
Continuing Use of HVA and Irrigation Practices. Same as above, but focusing on changes in these practices relative to the initial follow-up years.	Last Agricultural Season
Agricultural Production. Total amount of specific crops grown; amount of crops grown per square meter; total value of all crops cultivated.	Last Agricultural Season
Livestock. Number of cows, pigs, and sheep owned.	As of Survey Date
Revenue from Agricultural Production. Value of crops sold; total value of all crops (including those sold, bartered, or consumed).	Last Agricultural Season
Agricultural Costs. Expenditures on fertilizers, pesticides, irrigation water, hired labor, rented equipment, and taxes (individually and in total).	Last Agricultural Season
Profit from Agricultural Production. Revenues minus costs—the income from agricultural activities.	Last Agricultural Season
Income from Employment. Whether household head, spouse, and any grown children were employed (besides work on the family farm); total earnings from employment.	Last Month
Income from Pensions, Remittances, or Social Programs. Can also be added to profits and employment income to construct a rough measure of total income.	Last Month
Household Consumption. Expenditure on purchased food, health care, housing products, utilities, and transportation; cost of purchased goods (converted from monthly to annual) plus value of crops consumed by the household.	Last Month/Last Year

Survey Nonresponse. All interviews are conducted in person, and AREG devotes just one day for interviews in the majority of villages. Therefore, survey nonresponse is a concern. Substantial survey nonresponse can damage the validity of impact estimates. Nonresponse weights can account for some of the differential nonresponse, but only to the extent that nonresponse is explained by household characteristics that are known for both respondents and nonrespondents. In the worst case, survey nonresponse might be different for treatment and control villages, contradicting the core assumptions of a random assignment design. More commonly, however, survey nonresponse affects both the treatment and control groups equally. In this scenario, the impact estimates remain *internally* valid, but may not generalize beyond the select group of survey respondents. As a salient example, if farmers' absences are due to trips to the markets to sell their produce, then the respondents may have a disproportionate share of the less engaged farmers, for whom program impacts could be minimal.

We have instituted several safeguards against survey nonresponse. Working with mayors to clean the lists in advance can help in this regard. Whenever possible, village mayors would also contact the sampled households in advance to ensure they would be available for interviews on the day AREG visited their village. In instances where a household is not available on the first attempt, interviewers would return to the household throughout their time in that village. AREG also has reserve lists of farmers which they can draw on to help them reach their targets for completed interviews within each village. Nonetheless, AREG's methods are (understandably) designed more for accomplishing the targeted number of completed interviews within the short interview period than for ensuring high response rates. This will be an important issue to revisit with the survey contractor in the first follow-up survey next year.

Follow-up Surveys. Ideally, each round of the FPS would interview the same set of households, yielding a longitudinal data set. Analytically, longitudinal data allow for the cleanest estimation of program impacts, and also provide the most statistical precision, because changes from the baseline to the follow-up period are not confounded with sampling variability. As a practical matter, however, it may not be as easy to track specific households from year to year. Our plan is to survey the same set of households in subsequent rounds of the FPS to the extent possible, but given the nonresponse issues described previously, we anticipate that these will need to be supplemented with additional households, yielding a mixed longitudinal-repeated cross sectional data set. The sample frame will remain consistent, so as to avoid having the samples for treatment and control villages diverge over time.

IV. IRRIGATION PIU DATA

The Irrigation PIU has an impressive database for a subset of the WUAs, and by the end of the Compact, it is planned that information on members of all WUAs will be included in their databases. These data can possibly be used to supplement the survey data in two important ways. First, they provide some outcome measures that would not be obtainable from farmers, such as energy use and water distribution. Second, they will provide data on some outcome measures for the entire population of registered WUA members in Armenia. These data items are defined in more general terms than the survey data—for example, the amount of land the WUA member plans to grow wheat on, but not actual production—but they can still be used to inform us about broad national and regional trends. (The analyses of this data will likely be conducted by a local Armenian contractor, with oversight from MPR as needed.)

V. ESTIMATING PROGRAM IMPACTS

Random assignment ensures that, on average, treatment group villages and control group villages are the same, with the exception that treatment group villages are offered WtM training. Hence, the difference between the mean of the outcome of interest for the treatment group and the mean for the control group yields an unbiased estimate of the WtM program's impact. The precision of the impact estimates can be improved, however, by controlling for other covariates in a regression model. Regression adjustment can also help alleviate any differences between the treatment and control groups in baseline characteristics that arose by chance.

Core Specification. The survey data will be cross-sectional, with a new cross-section of respondents drawn each year.[1] Given this data structure, our econometric specification is designed to compare how treatment group villages changed over time to how control group villages changed over time, controlling for idiosyncratic differences in the two groups. The basic model can be expressed as follows:

(1) $y_{ivt} = \beta' x_{iv} + \lambda T_v \times F_t + \theta F_t + \eta_v + \varepsilon_{ivt}$,

where y_{ivt} is the outcome of interest for household i in village v at time t (where $F_t = 0$ in the baseline year and 1 in the follow-up year); x_{iv} is a vector of time-invariant characteristics of household i in village v; t accounts for any time trends between the base and follow-up years; T_v is an indicator equal to one if village v is in the treatment group and zero if it is in the control group; η_v is a village-specific error term (a village "random effect"); and ε_{ivt} is a random error term for household i in village v observed at time t. The parameter estimate for λ is the estimated impact of the program.

The vector of baseline characteristics x_{iv} will include both household and village characteristics. At a minimum, we will control for village characteristics such as the geographic region, WUA, and the baseline water conditions. We will also control for household size and composition, and characteristics of the household head, namely, education level, gender, age, and number of years farming. In the framework of a repeated cross-sectional model, however, the characteristics that are included must be restricted to those that are unaffected by the WtM programs. We must be careful with land holdings, for example, as the WtM program could conceivably induce some farmers to cultivate more land, and controlling for it would therefore understate the full program impact.

---

[1] As described previously, there will be substantial overlap in the household samples from the baseline year and subsequent years, but the samples will likely not be identical. If, however, the survey in subsequent years uses the same sample, we will be able to employ panel (longitudinal) data models. The intuitive interpretation of panel data models is similar to models of repeated cross-sectional data, but the estimation techniques differ somewhat from those described here.

The model in equation (1) is designed to answer the general research question, "How have villages in treatment group changed from the baseline year to the follow-up year, relative to villages in the control group?" This core model can be tweaked in a variety of ways to explore alternative specifications. A simple example would be to allow the time trends to vary across regions. The specification also (implicitly) weights all respondents equally, which could be modified to either give all villages equal weight, or weights equal to the village populations.

Such explorations would not change the general interpretation of the impact estimate, but they can provide insights on two important issues. First, and of most direct interest, we can explore how robust the impact estimates are to these alternative specifications. Beyond this, however, the other regression covariates may be of independent interest, and may also provide context for interpreting the impact estimates.

Pooled Model. Instead of using data from only the base year and one follow-up year, we can also pool data from multiple waves of follow-up year surveys. The econometric specification would be very similar to (1), but with a separate impact estimate for each of the *n* follow-up years:

$$(2) \quad y_{ivt} = \beta' x_{iv} + \lambda_1 T_v \times F_{1t} + \lambda_2 T_v \times F_{2t} + \ldots + \lambda_n T_v \times F_{nt} + \theta_1 F_{1t} + \theta_2 F_{2t} + \ldots + \theta_n F_{nt} + \eta_v + \varepsilon_{ivt},$$

where $F_{nt} = 1$ if $t = n$ and 0 otherwise.

These impact estimates can then be compared to one another to see how program impacts changed over time, and could be particularly important to see whether any impacts on farming practices that are observed early on persist, and also whether impacts on longer-term outcomes, such as agricultural productivity, grow after farmers have had more time to implement new techniques and benefit from their innovations.

Clustering. The estimation techniques must take into account the correlation of outcomes for households in the same village, as they may be exposed to similar idiosyncratic influences that are not otherwise captured in the regression model, and therefore, the individual households cannot be considered statistically independent. As an example, a particular village might have abnormally good or bad weather, or could experience other economic "shocks" that are unrelated to the training program but nonetheless affect the entire village. The econometric models will account for this clustering with methods that allow flexibility in the correlation structure of the error terms.

Impact on Participants Only. Randomly assigning communities to be eligible for WtM training programs provides an unbiased estimate of the impact of offering this training in the villages selected for training—the "intent to treat" (ITT) effect. The ITT effect combines the effect of the intervention on both participants and non-participants in treatment villages. In many contexts, people who are offered program services but opt out of participating are unaffected by the program, while in other situations the program may nonetheless have within-village spillover effects on the outcomes of non-participants. By including questions about both participation in WtM training and adoption of WtM techniques, we will be able to determine whether there are sizable within-village spillover effects present, and how best to account for them.

When spillover effects are known to be minimal, a simple but powerful adjustment can be made to calculate the effect of the training program on participants—the effect of "treatment on the treated" (TOT). This adjustment—known colloquially as the Bloom adjustment—calculates the effect of the training program on participants by dividing the estimated impact (the ITT) by the participation rate. The intuition for this elegant result is that, if the effect of the program on non-participants is known to be zero, the estimated impact can be attributed entirely to the proportion of the treatment group that actually participated in training. Importantly, however, while the Bloom adjustment can potentially be used to account for non-participation in the impact estimate, it cannot alleviate the problem non-participation introduces for the *variance* of the impact estimate. If participation rates are low, we will not be able to detect impacts that are statistically reliable.

Tables. The reports will include a variety of tables and figures with descriptive statistics on the data; however, the focus will be on estimates of program impacts and their statistical significance. We will also report regression-adjusted means for the treatment and control villages—that is, the means for the two groups if they had identical village and household characteristics. We will also report estimates of the TOT effect. Table 4 provides an example of the structure of these tables.

Table 4. Example of Impact Estimates Table

	Treatment Mean	*Control Mean*	*Program Impact*	*Impact on Participants*	*p-Value of Impact*
Total Value of Crops					
Total Ag. Expenditures					
Profit from Agriculture					
Etc.					

Notes: */**/*** indicate statistical significance at the .10/.05/.01 level. Treatment and control means are regression-adjusted to account for idiosyncratic differences in village or household characteristics of the two groups. Impact on participants is calculated using the Bloom adjustment, which divides the program impact by the participation rate.

Subgroup Analysis. For many of the outcome measures, it is conceivable that the effects of the interventions will vary by observable characteristics. Estimating differential impacts on female-headed households, for example, is of particular interest to MCC. We will examine whether the interventions' effects differ for key subgroups defined by the characteristics of the households such as gender, age, and level of education of the household head; size of the household; or size of farm holdings operated by the household. Similarly, we will also examine how effects vary by subgroups defined by village characteristics.

It is straightforward to embed subgroup estimates into the framework of equation (1). To do so, we include an interaction term that distinguishes treatment group members in subgroup *S* from those who are not in the subgroup:

(3) $$y_{ivt} = \beta' x_{iv} + \lambda_{S=1} T_v \times F_t \times (S_{iv} = 1) + \lambda_{S=0} T_v \times F_t \times (S_{iv} = 0) + \theta F_t + \eta_v + \varepsilon_{ivt}$$

In equation (3), the estimate of $\lambda_{S=1}$ represents the estimated impact for members of subgroup *S*, and we can test whether the impacts differ for members of that subgroup compared to everyone else by statistically testing whether $\lambda_{S=1}$ and $\lambda_{S=0}$ are equal.

Distributional Effects. The implicit focus of the analysis plan outlined above is on examining differences on the mean household. In conducting the analysis, it is also important to examine whether the interventions' effects vary at different levels of the outcome distribution. For example, the impact on agricultural real income for households with very low or very high income may differ from the impact on households at the mean. Specifically, the training programs may be such that only the higher-income households will benefit, if, for example, implementing the techniques taught in training requires investment in equipment that lower-income households cannot afford. Conversely, the poor in the community might benefit more than the wealthy if the training focuses primarily on techniques that are useful only to smaller-scale farms.

As Armenia has among the highest levels of income inequality in Europe, this distinction is not a trivial one. We will use quantile regression analysis to determine whether the intervention effects vary at different points in the distribution. Quantile regressions are analytically appealing because, similar to standard regression analysis, the quantile regression coefficients have direct and simple interpretation, thereby making it very appropriate for communicating impact estimates with policymakers.

Estimating impacts for specified quantiles starts with the same regression model as a standard model. The difference is in the methodology for estimating the parameters, which in turn, affects the interpretation of those impact estimates. While a standard regression model compares the impact for mean households, a quantile regression instead compares the impact of the interventions for a specified percentile, such as the 25th or the 75th percentile. Quantile regressions at the 50th percentile, the median, are also more robust to the influence of extreme outliers in the data, and thus can serve to validate the findings from standard regression analysis.

V. PLANS FOR REPORTING

We will submit reports on the WtM activities at three points in the lifecycle of the program. The first report will cover the baseline FPS, and will be a short report focusing on the current state of the villages in the evaluation. We will submit a draft of this in Spring of 2008. The second report will cover the second round of the FPS, after the training programs have begun at least one round of training in most villages. This report will focus on the intermediate outcomes, to gauge participation rates and preliminary adoption rates for the new technology and practices. The final report will follow the fourth year of the Compact, the last year before the control group villages will become eligible for WtM training. This report will focus on the longer-term outcomes, but as discussed, it will also examine intermediate outcomes such that we can assess not only whether there have been tangible impacts on poverty and household income, but also whether there is evidence from the intermediate outcomes that the full economic impact of the WtM may not have been fully manifested yet.