



December 11, 2015

MEMORANDUM

To: Moldova Evaluation Management Committee Members: Sixto Aquino, Jack Molyneaux, Kristin Penn, Jeremy Streatfeild, Leslie McCuaig

Cc: MCA-Moldova M&E, Andrei Bat  
USAID Moldova, Rodica Miron

From: Rebecca Goldsmith, M&E

Re: Cancellation of the Moldova Growing High Value Sales Activity Value Chain Training Impact Evaluation

The purpose of this memo is to document the cancellation of the Moldova Growing High Value Sales Activity Value Chain Training Impact Evaluation. The MCC Moldova Team believes that very limited learning would be achieved by completing the evaluation and that it does not warrant the remaining cost. This decision was made after in-depth analysis of the benefits and costs of the impact evaluation by MCC along with Mathematica Policy Research (Mathematica; the evaluator), MCA-Moldova, United States Agency for International Development (USAID), and DAI (the implementer).

The Moldova Compact's Transition to High Value Agriculture Project (THVA Project) was designed with four complementary activities to increase the ability and willingness of farmers to make the transition to higher value fruit and vegetable production. The THVA Project's activities were expected to: (i) rehabilitate up to 11 irrigation systems through the Centralized Irrigation System Rehabilitation Activity (CISRA); (ii) provide technical assistance and capacity building to support legal transfer of management and operations of MCC-rehabilitated systems from the Government to Water User Associations through the Irrigation Sector Reform Activity (ISRA); (iii) provide term financing and technical assistance to support high value agriculture-related investments by farmers and rural entrepreneurs through the Access to Agricultural Finance Activity (AAF); and (iv) provide market development support and technical assistance and training to help producers and agribusinesses better access high value agriculture markets and support the shift to high value agriculture at the production and postharvest level through the



Growing High Value Agriculture Sales Activity (GHS). The GHS Activity was undertaken jointly with, and administered by, USAID.<sup>1</sup>

The GHS Activity included the following four Sub-Activities:

- Develop and Expand Market Opportunities for Moldovan High Value Agriculture Sub-Activity
- Training to Upgrade Production and Meet Buyer Requirements Sub-Activity
- Demand-Driven Technical Assistance to Upgrade the High Value Agriculture Value Chain Sub-Activity
- Implement Recommendations for an Improved Enabling Environment Sub-Activity

The evaluation that is the subject of this memo was designed to evaluate the second sub-activity of the GHS Activity: Training to Upgrade Production and Meet Buyer Requirements.<sup>2</sup> The final evaluation design approved in January 2012 stated that the impact evaluation would answer these research questions:

1. What is the impact of GHS farmer training on adoption of new practices, production, sales, and farm income within the context of a value chain project? Do these impacts vary by value chain?
2. Does distance from a GHS farmer training site affect participation in GHS farmer training?
3. To what degree are new practices adopted by value chain participants who do not themselves participate in GHS farmer-training activities? Can adoption by nonparticipants be attributed to program ripple effects, rather than broader trends?
4. Is the economic rate of return (ERR) for the GHS training Sub-Activity large enough to justify the investment?

The locations that could be used for training activities were randomly determined from among a set of potential training sites to create a treatment and control group for the evaluation. After analyzing baseline data and training participation rates, Mathematica reassessed the costs and benefits of the evaluation, and recommended cancelling the evaluation. A few key challenges were cited by Mathematica as major limitations to what could be learned from the evaluation

---

<sup>1</sup> The GHS Activity was managed by USAID as part of the Agriculture Competitiveness and Enterprise Development (ACED) program.

<sup>2</sup> The evaluation was called the GHS Value Chain Training Impact Evaluation because the training focused on farmers already working in certain value chains. This was meant to differentiate it from the CIS Training which was a different training program targeting different farmers under the same GHS Activity.

(The full memorandum from Mathematica is attached and gives detailed explanations to the following points).

1. Most importantly, the statistical power of the evaluation was lower than expected because the evaluation sample is smaller than expected and the share of sampled farmers that were trained is lower than expected, reducing the ability of the evaluation to detect expected impacts. For example, with the anticipated training rate in trained communities/value chains in the treatment group (of 30 percent or less), the evaluation would only be able to detect impacts if the percentage of farmers using a specific practice or the average number of practices used more than doubles relative to the control group.
2. Additionally, there are a number of factors that may lead to inaccurate reporting. For instance, because agricultural trainings are commonly held in both treatment and control areas, it is difficult for participants to accurately report whether the training they attended was a GHS training or a training sponsored by a different project/organization. Also, as learned from the baseline survey, simply asking farmers whether they used a specific type of practice may lead to inaccurate reports due to the fact that some of the practices covered in trainings have fairly complex definitions.
3. Lastly, despite the fact that random assignment was stratified on targeted crop, there is a potentially important crop imbalance between the treatment and control groups due to differences in the number of farmers per community.

In addition to the aforementioned reasons, the costs of the evaluation are very high relative to the costs of the intervention. If the evaluation is continued, considerable outstanding costs for data collection, analysis, and reporting would be incurred.

Furthermore, the GHS Activities will be evaluated as part of the THVA Project Evaluation. The THVA Project Evaluation design was approved in October 2015 by MCC and will provide a comprehensive evaluation of the whole project including the GHS Activities. The THVA Project evaluation is designed to answer these research questions through a combination of impact and performance evaluations:

1. Were the expected results realized from the THVA program logic (with priority on the medium-term outcomes)? For example, to what extent did hectares of irrigated crops, hectares under intensive and non-intensive high-value agriculture, prices, and sales increase in the Central Irrigation System and border areas?

2. If results were not realized, why not?
3. What was the contribution of each activity/sub-activity to the results that were realized (this includes analysis of each sub-activity for ISRA, CISRA, GHS, and AAF)?
4. How did THVA affect land ownership, leasing, and land values in the CIS and border areas?
5. How are the results from the project distributed?
6. Are there indications that some of the long-term outcomes will be realized?
7. What lessons can be drawn from analysis of the design, implementation, and results of the THVA Project?
8. What is the ex post ERR of the THVA Project?

The third research question above will attempt to analyze the contribution of GHS training to the observed agricultural outcomes. This is similar to the intent of the first research question that the GHS Value Chain Training Impact Evaluation was designed to answer, but it will be conducted qualitatively rather than quantitatively as was originally planned. The other three impact evaluation questions about the value chain training will not be answered by the THVA Project evaluation.

In sum, the team has come to the conclusion that the learning that the evaluation could provide is extremely limited and does not outweigh the costs and the Moldova Growing High Value Sales Activity Value Chain Training Impact Evaluation will be cancelled. The evaluation design report, baseline report, and this cancellation memo will be made available publicly on MCC's [Evaluation Catalog](#).

Sincerely,  
Rebecca Goldsmith  
Director  
Monitoring & Evaluation



## **Attachments**

(1) Benefit-cost assessment Memorandum by Mathematica – 12/14/2014

---

**MEMORANDUM****TO:** Rebecca Goldsmith**FROM:** Jane Fortson and Evan Borkum**DATE:** 12/14/2014  
MDOVA-128**SUBJECT:** Benefit-cost assessment of the GHS value chain training evaluation

---

The Growing High-Value Agriculture Sales (GHS) activity is one of several activities being funded in Moldova by MCC under the Transition to High-Value Agriculture (THVA) project. It consists of four complementary subactivities: (1) high-value agriculture-market development and expansion (including end-market studies and linkages to potential investors); (2) training for producers to upgrade production and meet buyer requirements; (3) demand-driven technical assistance to enterprises, associations, and cooperatives; and (4) the improvement of an enabling environment for high-value agriculture (including strengthening phytosanitary inspection and testing capacity).

Mathematica Policy Research is conducting an impact evaluation of subactivity (2), the GHS value chain training subactivity, which provides targeted training on cultivation and post-harvest practices to existing producers of high-value crops in specific value chains (for example, apples, plums, and tomatoes). These producer trainings aim to help farmers of high-value agriculture products upgrade production and improve the efficiency of post-harvest activities, such as processing, transporting, and delivering products to consumers. Trainings target farmers who are already cultivating specific high-value crops, and are delivered using classroom instruction, demonstration plots, farmer field days, and other methods. To evaluate this subactivity, Mathematica is using a random assignment design in which potential training sites are randomly assigned to treatment and control groups. Trainings in the treatment sites began in late 2011 and are currently close to being completed.

As the GHS value chain training implementation and evaluation have proceeded, several challenges have been identified that potentially limit what can be learned from the evaluation. In March 2013, MCC prepared a memo summarizing these limitations and detailing the costs and benefits of the evaluation. Based on this analysis, it was determined that the evaluation should proceed, with a focus on practice use as the primary outcome of interest (statistical power for longer-term outcomes such as production, sales, and income was insufficient to detect meaningful impacts).

The purposes of this memo are: (1) to reassess the costs and benefits of the GHS value chain training evaluation in light of new information that has emerged as the evaluation has proceeded, and (2) to address requests by DAI, the GHS activity implementer, to conduct training in some of the control communities. Overall, our analysis suggests that the benefits of the evaluation are likely

MEMO TO: Rebecca Goldsmith  
FROM: Jane Fortson and Evan Borkum  
DATE: 12/14/2014  
PAGE: 2

to be very limited and do not justify the costs; we therefore recommend that the impact evaluation be canceled in its current form. The GHS value chain training subactivity would still be evaluated, though less rigorously, through Mathematica's overarching evaluation of the THVA project. Given this recommendation, DAI's request to conduct trainings in some control communities is somewhat moot; nevertheless, we find that approving the request is unlikely to substantially harm the evaluation should it continue. Below, we provide a brief overview of the current evaluation design, reassess the benefits and costs of the evaluation with a focus on new challenges that have arisen, consider the implications of DAI's request to conduct training in some of the control communities, and explain why we believe the evaluation in its current form should be canceled.

## **1. Overview of evaluation design**

The impact evaluation of the GHS value chain training subactivity uses a random assignment design. At the outset of the evaluation, DAI identified 80 potential training communities for inclusion in the impact evaluation. In each community, DAI identified one or more value chains ("targeted value chains"), with the expectation that, if a community were selected for training, trainings in that community would focus on those value chains.

In September 2011, potential training sites were randomly assigned to treatment and control groups (48 treatment, 32 control). Training activities could be conducted at sites assigned to the treatment group, and would not be conducted at (or near) sites assigned to the control group. To ensure representation of all value chains and regions in both the treatment and control groups, we conducted random assignment within strata defined by targeted value chain (or combinations of targeted value chains) and region. The sample for the evaluation consists of a representative sample of farmers in the targeted value chains in each community. A baseline survey for the evaluation was conducted in early 2013, with a follow-up survey to estimate impacts planned for late 2018 or early 2019.

The evaluation design must also account for the fact that some farmers who live in treatment communities may choose not to participate in training, while some farmers who live in control communities could choose to travel to attend training. Hence, random assignment does not necessarily separate farmers who attend training from those who do not. Instead, random assignment changes the *probability* that farmers attend training, assuming that those who live in treatment communities are more likely to attend. The evaluation can account for this feature in estimating the impacts of training participation; however, for the evaluation to be viable, farmers in treatment communities must be substantially more likely to participate in training than those in control communities. If training participation rates are not sufficiently different between farmers in the treatment and control communities, the evaluation will not be able to detect meaningful impacts.

MEMO TO: Rebecca Goldsmith  
FROM: Jane Fortson and Evan Borkum  
DATE: 12/14/2014  
PAGE: 3

## 2. Reassessment of costs and benefits of the evaluation

The primary expected benefit of the evaluation was that it would provide rigorous evidence on the impact of producer training on practice use, on which there is limited existing evidence in the literature. Because the GHS value chain training is being implemented in conjunction with several other subactivities designed to address barriers to the sale of high-value crops, the impacts of training might be expected to be higher than the impacts of training in other contexts. The evaluation could be informative for future MCC investments in the agricultural sector, as well as for policymakers, implementing agencies, and other stakeholders in Moldova and other countries.

However, as the evaluation has proceeded, additional challenges have emerged that severely limit the evaluation's ability to provide this evidence:

- **Statistical power is lower than expected.** The statistical power of the evaluation depends on two factors. First, it depends on the number of communities and farmers in the sample (because the design involved random assignment of communities, the number of communities is particularly important). Our original assessment of statistical power assumed that DAI would conduct trainings in all 48 treatment communities (in all targeted value chains) and in none of the 32 control communities. However, in practice, DAI did not conduct trainings in all treatment communities (and in some cases trainings were conducted, but not in the targeted value chains from which we drew our sample). Overall, roughly one-third of our weighted treatment sample is in a community and value chain in which DAI did not provide trainings. DAI also conducted trainings potentially affecting most of the sample in one large control community.<sup>1</sup> In addition, several communities had no farmers in the targeted value chains, and will not contribute to the evaluation.

Because not all treatment communities were trained, some control communities were trained, and some communities had no farmers in targeted value chains, we also have to drop additional communities in some random assignment strata, further decreasing the sample size. This is because each stratum must include both treatment and control farmers in order to contribute to the evaluation. After accounting for these changes, some of the relevant strata are left with very few (if any) control or treatment farmers, making it necessary to drop the entire stratum. Unless DAI conducts more trainings in

---

<sup>1</sup> This large control community was trained prior to random assignment, but we only discovered that it had been trained after random assignment had been conducted; as a result, we approved DAI's request to conduct additional trainings in that community. DAI also conducted training in a second control community, though we do not recall a request being made. This community had few sampled farmers in the value chains in which DAI conducted training. Because most of the sample in this community is in untrained value chains, it will still contribute to the evaluation.



MEMO TO: Rebecca Goldsmith  
FROM: Jane Fortson and Evan Borkum  
DATE: 12/14/2014  
PAGE: 4

the remaining treatment communities, the evaluation would functionally use data from only about 52 communities, 28 in the treatment group and 24 in the control group.

Second, statistical power depends critically on the difference in GHS value chain training participation between farmers in treatment and control communities. Our current assessment—based on self-reported training attendance and earlier attempts to match the names of sampled farmers to lists of trained farmers—is that the participation rate in treatment communities is likely to be low, even in communities in which trainings were held in targeted value chains. In part, this is because training participants were frequently members of the same farm, while each farm only contributes one observation to our sample.<sup>2</sup> In addition, there is substantial heterogeneity in the number of trainings across communities, with some communities receiving only one training and others receiving multiple trainings. Therefore, the share of treatment farmers attending more than one training—who are most likely to experience large impacts—is likely to be even lower.

As a result of both the decreased sample size and low training rate in the treatment sample, the ability of the evaluation to detect expected impacts—even on practice use—is extremely limited. For example, with the anticipated training rate in trained communities/value chains in the treatment group (of 30 percent or less), we would only be able to detect impacts if the percentage of farmers using a specific practice or the average number of practices used more than doubles relative to the control group.

- **Identifying training participation is challenging.** Agricultural trainings are commonly held in both treatment and control areas. Because GHS value chain trainings are typically contracted out to regional training providers who also conduct other trainings, it is difficult for participants to accurately report whether the training they attended was a GHS training or a training sponsored by a different project/organization. Matching reported characteristics of trainings attended to those of GHS value chain trainings (for example, based on date and value chain) is also problematic because respondents may have difficulty recalling this information. In theory, the most accurate way to identify training participants is to match their names directly with the list of trainees maintained by DAI. However, we have found this exercise to be fraught with error because of differences in spellings of names and many individuals having similar names. It is also possible that the list of trainees maintained by DAI does not capture all farmers who participated in training (if training providers did not keep accurate records).

---

<sup>2</sup> In control communities in which trainings were not conducted, training participation rates are low—as required by the design. The main challenge is low participation rates among the sample in treatment communities.

MEMO TO: Rebecca Goldsmith  
FROM: Jane Fortson and Evan Borkum  
DATE: 12/14/2014  
PAGE: 5

- **It is difficult to accurately measure practice use.** Use of practices taught in value chain trainings is the primary outcome of interest for the evaluation. However, one lesson from the baseline survey is that simply asking farmers whether they used a specific type of practice may lead to inaccurate reports. In particular, some of the practices covered in trainings have a fairly complex definition, and farmers might report that they used the practice without fully understanding what it entails. If training increases treatment farmers' familiarity with these practices and their correct definition, training could even *reduce* reported practice use in the treatment group relative to the control group. Obtaining more accurate information about practice use would require substantially increasing the complexity of the survey, which is already very lengthy.

In addition, given the large number of training practices, it is not clear what measure of practice use to focus on in the impact analysis. We have discovered that training varies considerably both across communities and by value chain, so that not all practices are covered in all trainings. However, the list of practices captured in our survey is a subset of all training practices (which we asked DAI to prioritize to make the survey length manageable). Therefore, even if farmers do attend trainings, they might not learn about all the practices that we are measuring. As a result, impacts on the use of specific practices measured in our survey—one approach to measuring practice use—might be limited (moreover, examining impacts on so many practices is likely to result in many spuriously significant impacts). Another possible approach is to use a summary measure such as the total number of practices used (of those captured in our survey), but that is a fairly crude measure that may be difficult to interpret, and is also limited because the survey does not cover all practices.

- **There is a potentially important crop imbalance between the treatment and control groups.** Our baseline analysis suggested that the treatment and control samples were broadly similar at baseline in most key characteristics and outcomes, but that there were some important differences in the targeted crops cultivated by the two samples (despite the fact that they were randomly assigned). Specifically, farmers in the treatment sample were substantially more likely to cultivate targeted table grapes and substantially less likely to cultivate targeted vegetables relative to the control sample. Although neither of these differences is statistically significant, the magnitudes are relatively large and could be a cause for concern if production or market shocks unrelated to value chain training affect specific crops. The primary cause of these differences was substantial variation in the number of farmers per community cultivating specific targeted crops. Random assignment sometimes led—by chance—to an imbalance in the number of eligible farmers in the treatment and control groups; as a result, there are sometimes large differences in cultivation of targeted crops. It will therefore be important to consider differences in cultivation of targeted crops when interpreting the impact evaluation results, and to explore the extent to which these compositional differences might be driving the results. However, our ability to do so

MEMO TO: Rebecca Goldsmith  
FROM: Jane Fortson and Evan Borkum  
DATE: 12/14/2014  
PAGE: 6

will be limited because we have limited statistical power to estimate impacts by value chain, given the smaller sample sizes for each individual value-chain.

At the same time, the costs of the evaluation are very high relative to the costs of the intervention. If the evaluation is continued, considerable outstanding costs for data collection, analysis, and reporting will be incurred. The evaluation also places constraints on implementation (to maintain the integrity of the control group), which have the potential to impose implementation costs by reducing program efficacy. Given the implementation timeline, canceling the evaluation now would be unlikely to have a substantial effect on implementation (implementation in control communities would be limited in any case), but some implementation costs may be incurred.

### **3. DAI's request to conduct trainings in control communities**

DAI has requested permission to conduct trainings in additional control communities where they believe trainings will have large potential impacts. Some of these trainings will be conducted in areas that will benefit from the irrigation-related activities of the THVA project and will be targeted at non-high value agriculture farmers (to help them transition to high-value agriculture). The topics for these trainings include a small number of practices that we are measuring as part of the value chain training evaluation, but most of the topics will be different (they will be more relevant to non-high value agriculture farmers). Because our sample for the GHS value chain training evaluation consists of high-value agriculture farmers, these CIS area trainings are unlikely to adversely affect the evaluation as long as they do indeed target non-high value agriculture farmers and as long as sharing of information with the farmers in our sample is limited. DAI has also requested permission to conduct value chain trainings in (or near) two additional control communities. Because our sample includes very few (if any) farmers in the relevant value chains in these communities, this would also not pose a major challenge for the evaluation, although there is always a risk of additional requests for training roll-out to control areas. *Therefore, DAI's request to conduct trainings in these control communities (CIS area training in CIS Cosnita and CIS Grozesti; value chain training in Stefan Voda and Negureni) does not threaten the evaluation and we recommend that the request be approved.*

MEMO TO: Rebecca Goldsmith  
FROM: Jane Fortson and Evan Borkum  
DATE: 12/14/2014  
PAGE: 7

#### **4. Conclusion and implications for the evaluation**

Overall, our assessment suggests that the benefits of the GHS value chain training evaluation are very limited, and do not justify the evaluation costs. Given the reduction in statistical power, we would not be able to identify the impacts of the subactivity, even if the other challenges—which are also severe—could be overcome. It is also unclear how the existing evaluation design or sample could be modified to address these challenges. For example, to compensate for dropped communities, we could increase our sample size of farmers in the remaining communities—but this would only lead to small improvements in statistical power given the clustered design and the low training rate.

We therefore recommend canceling the GHS value chain training evaluation in its current form. Instead, the evaluation of this subactivity could be incorporated into the broader performance evaluation of the THVA project, which is being developed. Although this evaluation will not be able to provide rigorous impact estimates, it will be better able to capture the complementarities between the value chain training subactivity and other GHS subactivities that are key to the THVA program logic.

cc: Ken Fortson; Candace Miller; Seth Morgan; Anca Dumitrescu; Irina Cheban;  
Bethany Simard; Alexander Johann