

## **Summary of the Impact Evaluation**

The Water-to-Market activity in Armenia was actually composed of several components, and to structure my review, it is worth first summarizing those components. The initial impact evaluation was designed to measure the impacts of two training components, known as On-Farm Water Management (OFWM) training and the High Value Agriculture (HVA) training. After the randomized evaluation of those activities was designed, Mathematica Policy Research was also asked to conduct evaluations of three other components of the Water-to-Market activity: the Access to Credit subactivity, the Institutional Strengthening sub-activity (ISSA), and the Post-Harvest, Processing, and Marketing (PPM) subactivity. The evaluations of the latter three components are therefore either qualitative, or use the data collected to speak towards the actual impacts.

## **Impacts of the Water-to-Market Activity**

### *The OFWM and HVA Training*

The two training activities, the OFWM and the HVA trainings, had something of a mixed record in the impact evaluation of them. On one hand, the trainings did not appear to change behavior much. On-farm water management practices did not appear to change; farmers continued to use simple techniques rather than more advanced techniques such as gated pipes or drip irrigation. One of the main problems, noted in the report, is that incentives are not really aligned for farmers to save water through advanced techniques, particularly given that advanced techniques require investments. Water is billed based on the amount of land farmers have and the crops intended to be irrigated, rather than by the amount of water used, dampening the desire of farmers to save water; the only remaining incentive to save water is if yields then increase by using less water, increasing net returns to agriculture. An interesting follow-up experiment might be to change the water pricing scheme in a random set of villages trained towards the end of the activity, to observe whether or not farmers who know of improved irrigation techniques would then adopt them given a different pricing scheme. There is some evidence in the report that they might not, as farmers might not have enough access to credit to implement medium or advanced irrigation techniques.

Although there were no impacts on irrigation techniques, presumably through the HVA training there appear to have been some modest impacts on improved soil preparation and purchasing pesticides. However, there are no impacts (as modeled currently) on what farmers grow. Perhaps the most interesting impacts are in Table II.9., which lists the impacts on production in metric tons, land under cultivation, the revenues from crops sold, and the market value of the harvest. The market value of three high value crops increase by substantial amounts (tomato, vegetables, and potatoes), more than offsetting a decrease in the value of grapes harvested. Most of the gain in the total market value can be attributed to those three crops (there is also a

difference in the value of grain produced). In fact, the total agricultural profits (calculated in Table II.12) increase somewhat, by approximately 18 percent. Finally, in Table II.13 it is shown that there is substantial regional heterogeneity in the results: profits go up in mountainous regions, but not in the pre-mountainous regions nor the Ararat valley, where high value crops were already more prevalent. The changes in profits are curious, because there is no evidence of changes in behavior in other parts of the report.

This finding is quite intriguing, and it later also guides the economic rates of return (ERR) calculations, which show substantial returns (82 percent per annum if the impacts are sustained for 20 years). There are a few concerns I have with what we learn from these results, particularly as there is no evidence of behavior change leading up to the findings. I have a couple of explanations. First, it could be that outliers are driving the results, particularly as income tends to be a skewed variable; the authors recognize this fact, but do not try the obvious solution, which is to estimate impacts on the logarithm of profits, rather than the actual profits. The logarithmic operator is useful because it dampens the effect of outliers on average coefficient estimates. Although the report is quite opaque on distributions of variables, average profit levels seem high enough that negative variables would not be too much of an issue. Second, there seems to be information to exploit in the data to truly look at crop transitions. The data include measures of the crops grown in the baseline and endline; it seems that it would be worthwhile studying how the transitions actually take place, and linking those transitions to the crops for which average productivity changed. It could be that the HVA trainings in particular might have catalyzed specific types of crop transitions that are masked in Table II.11 because it does not attempt to control for the crop(s) initially being grown by the household.

Third, and somewhat tangentially, the computation of the ERR could be improved. What the authors currently do to get the average treatment effect on the treated is to take the intent-to-treat effect and then divide by the proportion of farmers who actually received the training. The proper way to do this calculation would rather be to estimate a regression of agricultural profits on an indicator variable for participation in training, using the treatment indicator as an instrumental variable. Since we know that the treatment indicator is a proper instrument, this calculation would give the per household impact of participating in training. Of course, it is likely to be as noisily estimated as the current estimate.

So, what can be concluded about the impacts of the training activities? First, the trainings themselves did not appear to change producer behavior much. There were some small changes in behavior identified that can be attributed to the HVA trainings, but little if anything should be attributed to the OFWM trainings. Yet there are some crops for which revenues increased; it would be worth understanding better whether there are specific transitions (from grains to tomatoes for example, even at the margin) that were clearly facilitated by the trainings. These transitions clearly would have taken place in mountainous regions, which had more to gain; these gains in the mountainous regions deserve further explanation. In general, the impacts we observe at this point are intriguing, but leave the reader asking for a deeper explanation.

### *Impacts of the Access to Credit subactivity*

The next component of the report attempts to measure impacts of the Access to Credit subactivity, in which a pool of \$8.5 million was made available to catalyze agricultural investments among farmers who attended WtM trainings. By 2011, this activity had given loans to over 1000 farmers. Unfortunately, only 27 loan recipients existed in both rounds of the data set. The credit subcomponent was very popular due to a low interest rate relative to other sources of agricultural credit; the ability to borrow in Armenian drams rather than dollars; and because of the relatively small size of loans allowed by the program. Evaluation in this case took place through difference-in-differences, using a portion of the sample that had not received WtM loans as the comparison group (some of those had received other forms of credit). There is clearly positive selection in determining which households received loans. Households had to fill out a detailed application and have clear plans for the funds. Not surprisingly, the household heads who have received credit appear to have higher education levels than others. Further, baseline characteristics suggest that those receiving credit were wealthier as well (Table III.7.). The modal farm household receiving credit uses it build a greenhouse.

Despite what appear to be somewhat large increases in household income among credit recipients, relative to all other households (rather than a good comparison group), differences in income between groups are not statistically significant. Therefore it is difficult to ascertain what the impacts were that are attributable to the credit activity itself. This is even more true as the credit recipients were positively selected, and therefore we consider that the estimates are likely biased upward.

The ERR calculations are almost surely biased upward because of the same positive selection of beneficiaries. That said, there is a return to the banks (in terms of collected interest) that is neglected in the calculation of the ERR as well.

Although we cannot say much about the precise impacts of the credit subactivity, its popularity appears to have filled a market niche. Not only is it popular, but repayment rates are high, meaning that it seems to have economic value to the banks distributing WtM credit. It will therefore serve three times as many beneficiaries as it had in 2011 by 2020. The fact that the credit is popular and can be paid back strongly suggests that the activity has a positive return to recipients. That said, it is not a particularly good anti-poverty device, since poorer households are less likely to receive it, but it does clearly play an important role in the economy.

### *Institutional Strengthening of Irrigation Management Entities (ISSA) Subactivity*

The ISSA subactivity was a \$4.9 million dollar project to build up the capacity of the 44 Water Users Organizations of Armenia. 8 of those organizations received more intensive help. On top of technical assistance, the organizations received office equipment and computer software. Finally, the ISSA attempted to influence irrigation policy through a policy reform that was taking place. A major issue is that the organizations are not self-sufficient; they need fairly substantial

government subsidies to operate. A goal was not to make these organizations self-sufficient, but rather to improve the way they collected funds and making them closer to self-sufficient.

The evaluation found that after milestones were set individually for all 44 organizations, 5 major milestones in terms of improvement were met in all of the organizations, and 4 other milestones were met by 40 of the 44 organizations. The donation of equipment, including computer equipment, was made conditional on meeting some of these milestones, which appears to have helped improve decision making. However, it is not clear how well these reforms translated to improved outcomes for the organizations, which really can be thought of as water delivery and cost recovery. The amount of water being delivered did not increase over time pre- and post-reform; we cannot measure how much additional water was saved rather than being wasted, which would be another good measure. The amount of fees collected improved somewhat, but the cost recovery rate had not reached the target of 53 percent for 2010.

It is difficult to assess what the overall impacts of the ISSA will be on water delivery in Armenia. On a positive note, through policy reform, better equipment, and more participatory decision making, it could be that the organizations have been put on a path towards requiring lower subsidies. On the other hand, the fact that leaders of the reform at the organizational level did not necessarily feel like a part of the reform could minimize its effectiveness later. Longer term monitoring might be necessary to learn the effectiveness of the ISSA in total.

#### *Post-Harvest Processing and Marketing Subactivity*

The fourth component of the WtM activity was the post-harvest planning and marketing (PPM) subactivity; it had a budget of \$4.2 million. The PPM included two primary components: a package of interventions meant to strengthen the value chain, which included instruction in post-harvest handling of crops and improved technologies for handling crops, including points with refrigeration; and a component that included a “buy Armenian” campaign. The former component also included establishing a market information system providing information about prices at 64 different markets around the country.

As with the previous two subactivities, there was no formal quantitative evaluation planned, due to the difficulty of constructing a control group. As a result, there is only evidence available from a qualitative evaluation of the subactivity, as well as administrative data.

The qualitative data, as well as reports from the field, have some positives. Farmers and others involved in trainings found the information on specific techniques quite useful. Also, according to data that were collected as part of the program there is some reason to believe it was successful. According to self-reports, many groups reported increases variables such as product quality, production, and identifying new markets.

On the other hand, there appear to be a number of drawbacks from the MPP as well. For example, it seems that the information about how to make market linkages was less useful for

farmers and groups; farmers still reported having difficulty selling their produce, even after improvements in drying techniques, etc. According to the implementing agency, the large quantitative targets for the number of groups to serve led to less contact per group. Further problems also existed; for example, farmers in Armenia are not inclined to work in groups (such as marketing cooperatives), because it seems like re-collectivization of a kind to them.

## **Lessons**

First, I want to focus on the impact evaluation of the trainings. One of the important messages in the impact evaluation report is that even though around 60 percent of farmers in treatment groups took part in the trainings, the qualitative information indicated that some farmers who participated in training were not actively farming, and in other villages mayors claimed that participants would get access to MCA credit, which was highly desired. One of the reasons that mayors might have suggested additional farmers participate in the training is that communities (or the implementing agency) were concerned about hitting the target number of participants. Clearly, there is a tension between target numbers of participants in projects and the quality of participation in projects in developing countries. One method of getting around the demand problem is to use what are being called “pull mechanisms,” in which actors within the economy upstream in the value chain “pull” rural residents to provide specific crops or use certain techniques. It is worth considering whether pull mechanisms might lead to better outcomes, although it is difficult to see how they could be used effectively in the context of OFWM trainings. One could envision offering contracts for high value crops conditional on the crops reaching a specific quality level, and then offering trainings on how to grow crops at that quality level; this is what is meant by a pull mechanism and might be worth trying in the future as an alternative method to ensure that farmers want to participate in trainings.

Second, I want to briefly discuss the credit intervention. The more I reflect on it, the more I think it was a success, although one has no “impact” evaluation that demonstrates the success—we need (as a profession) better measures for the success of something like revolving credit. From my perspective, the fact that it will continue to serve new farmers far into the future makes it quite successful, regardless of the fact that the ERR reported here is quite clearly positively biased as calculated.

Third, I want to discuss the PPM a bit more. It leads to some difficult questions—more and more emphasis is being placed on working through the value chain in agricultural interventions designed to improve productivity and income. These findings show it is possible to demonstrate some “progress” in improving value chains, using broad indicator variables such as “improved production.” Yet it is far more difficult to design interventions that will have a lasting impact. This point is even more salient if you consider how difficult it is to create a project with a positive rate of return. If, for example, the intervention had worked more intensively with half the number of groups, it is not clear how much more successful it would have been, but the cost per beneficiary would have been double. Again, this is a place where pull mechanisms could work (e.g. working with brokers who demand the product from

farmers), but the thinking about how to design pull mechanisms to make them effective still needs to be done.

Finally, I suggest that projects like the PPM might be better designed in small scale to work with one or two crops at first (in specific countries), to better develop the programs and to be able to construct impact evaluations around them. Such evaluations can help shed light on what works (perhaps testing a pull mechanism against a more standard project). When projects try to do too many things (e.g. work with too many crops), they lose focus, and it becomes more difficult to design evaluations of the projects as the mid-term objectives become less clear.