MEMORANDUM

TO: Rebecca Tunstall and Claudia Argueta

FROM: Larissa Campuzano and Randall Blair   **DATE:** 2/16/2011

ESVED-264

SUBJECT: Final Impact Evaluation Design for the Production and Business Services Activity of the Productive Development Project **- Revised**

EXECUTIVE SUMMARY

The Productive Development Project (PDP) is one of three large-scale projects financed under the 2006 Compact between the Millennium Challenge Corporation (MCC) and the Government of El Salvador.[1] The main objective of the PDP is to assist in the development of profitable and sustainable business ventures for poor individuals and organizations that benefit poor people in El Salvador's Northern Zone. The PDP comprises three activities: (1) Production and Business Services (PBS); (2) Investment Support; and (3) Financial Services. The final, core evaluation design for the PBS activity of the PDP is a *randomized rollout design*.[2] Mathematica adapted this design to each of the three value chains included in the impact evaluation: (1) handicrafts, (2) dairy, and (3) horticulture. To capture information on productive activities and the evaluation's key outcomes of employment generation and household income, Mathematica designed the Productive Development Surveys—or Encuestas de Desarrollo Productivo (EDPs)—which are tailored to the practices and goods of each of the three productive chains. In addition to employment and income information, the surveys will capture information on several intermediate outcomes, including production levels, sale prices, business practices, and technology adoption. Baseline EDP data collection took place between October 2009 and June 2010, and baseline results will be disseminated in March 2011. Follow-up data collection started in November 2010 and results from the first round of follow-up surveys will be available by

---

[1] The Compact's other two projects are the Human Development Project and the Connectivity Project.

[2] This memorandum describes the final impact evaluation design for the PBS activity of the Productive

[2] This memorandum describes the final impact evaluation design for the PBS activity of the Productive Development project, as agreed upon by Mathematica Policy Research, Millennium Challenge Corporation (MCC), FOMILENIO, and Chemonics. We describe our final impact evaluation design for the Investment Support activity in memo ESVED-238, submitted on July 1, 2010.

December 2011. Results from the second round of follow-up surveys will be available by January 2013.

A.   DESCRIPTION OF THE INTERVENTION

The PDP will use nearly $72 million in allocated funds to serve approximately 13,500 beneficiaries and create an estimated 11,000 full-time equivalent jobs.[3] The PDP comprises three activities: (1) Production and Business Services (PBS); (2) Investment Support; and (3) Financial Services. The PBS activity was designed to include pre-investment studies, training, and technical assistance to small farmers and business owners, in-kind contributions of agricultural and genetic materials, legal assistance, and other business development services. PBS targets producers in the following value chains: dairy, horticulture, fruit, handicrafts, tourism, and forestry. The Investment Support activity is designed to offer investment capital for competitively selected business proposals. Finally, the Financial Services activity provides technical assistance and financial resources to the banking sector and loan and output guarantees to small producers, as appropriate.[4] One overarching service provider, Chemonics, is coordinating and managing the various components of the PBS activity. In partnership with FOMILENIO, Banco Multisectorial de Inversiones (BMI) is coordinating the Investment Support and Financial Services activities.

The PBS activity was initiated with a pilot phase from 2008 to 2009. The objective of the pilot phase was to test processes and procedures for overall PBS implementation, generate lessons learned, and make the necessary changes to improve services. During this pilot phase, $5 million was disbursed to 13 productive projects. Projects included technical assistance for dairy farmers provided by TechnoServe, technical and material assistance to artisans provided by Aid to Artisans, and training provided by CARE to small farmers making investments in horticulture production. Pilot projects benefited 155 groups and 3,625 people with technical and material assistance.[5] No impact evaluation of the pilot phase of the PBS activity was completed.[6]

From 2008 to early 2010, the PBS activity focused on providing technical and material assistance to increase the production of small-scale producers. Services focused on increasing

---

[3] Plan de Trabajo de FOMILENIO, 2011.

[4] As of January 2011, the Investment Support activity had a total budget of $9.7 million, and the Financial Services activity had a total budget of $5.2 million.

[5] According to monitoring reports submitted by pilot phase implementers.

[6] However, a process analysis of the pilot was conducted.

and improving production through cheaper inputs, technology adoption, and enhanced business practices. For each value chain, one service provider covered all assistance to beneficiaries in the Northern Zone. The service provider for the handicrafts value chain, Aid to Artisans, worked strictly with groups of producers, whereas the other service providers (CLUSA for horticulture and TechoServe for dairy) worked with individual producers as well as organized groups. Due to changes in some service providers and contract renewals during late 2010, very few beneficiaries in any value chains received services from mid-September to late October 2010.

The PBS activity underwent a large-scale reorganization in late 2010. The second phase of PBS services (Phase II)—which formally began in September 2010—is fundamentally different from PBS services offered from 2008 to late 2010 (Phase I). Under Phase II, technical assistance for the horticulture chain will be delivered in tandem with assistance for the fruit value chain. In addition, the scope of technical assistance in the second phase is much larger: whereas assistance until late 2010 focused primarily on increasing production, current services are organized around producers' three basic needs: (1) productivity, (2) access to markets, and (3) business/marketing capacity. Organized by geographic regions, service providers in each chain work directly with beneficiaries to buy inputs in bulk, enhance production, and sell production in bulk. In addition, with the exception of TechnoServe in the dairy value chain, service providers in Phase II are different from those that provided technical assistance in Phase I. Swisscontact is the new provider for the handicraft value chain, and El Zamorano is the new provider for the horticulture value chain.

Modifications to the PBS activity under Phase II also include a commercial alliance between FOMILENIO, USAID, and Super Selectos, as well as the creation and/or strengthening of several producer-owned enterprises. Large-scale beneficiary-owned businesses will specify production levels and buy production from productive groups at a higher price. The producer-owned enterprise in the horticulture chain is *El Salvador Produce*, and the primary producer-owned enterprise in the dairy chain is *Lácteos de Morazán*. These enterprises will negotiate contracts with Super Selectos and other large buyers, and then organize and aggregate production through use of collection and supply centers. This includes 17 collection centers for the dairy chain, 10 workshops for the handicrafts chain, and 5 collection centers and 2 supply centers for the horticulture chain.

The PBS activity currently has several funding components: $21 million is allocated for general implementation (which includes $3 million to finance storage centers); $13 million is allocated for in-kind investments; $8 million is allocated to the *Fondo de Apoyo al Desarrollo Productivo* (FADEP), which will finance direct technical assistance to farmers related to production and sales, and provide $500,000 to strengthen the capacity of *El Salvador Produce*; and $4 million is allocated to the *Fondo de Iniciativas Productivas* (FIP), which will finance ongoing productive projects similar to pilot projects. The FIP has funded activities including a project to strengthen fruit production

systems in Cuscatlán y Cabañas ($341,000), a project to link dairy farmers in la Unión ($344,000), and a project linking producers of chipilín and loroco to markets in the United States ($670,000).[7]

MCC has contracted with Mathematica to design and conduct evaluations of the first two PDP activities—PBS and Investment Support. These two activities will be evaluated under different designs: the PBS evaluation will use the randomized rollout design outlined in this document, whereas the Investment Support evaluation will use a case study design (see memo ESVED-238, dated July 1, 2010). However, the nature of these services may allow for beneficiaries to receive a mix of the services offered under the PDP activities. To the extent possible, this impact evaluation will attempt to assess the separate effects of each activity. If separating the effects is not possible, the evaluation will assess the effects of the mix of services provided by both activities.

The rest of this memorandum describes the PBS evaluation design in detail, including the research questions it will address and the methods we propose for conducting it.

B.   KEY RESEARCH QUESTION

The impact evaluation addresses the following primary research question: What impact did the offer of PBS services by FOMILENIO/MCC have on employment creation and eligible producers' household incomes? Although assessing the impact on employment creation and eligible producers' income represents the main goal of the evaluation, we will also address impacts on intermediate outcomes, such as production levels, sale prices, business practice adoption,[8] technology adoption,[9] product diversification,[10] and value chain integration.[11]

---

[7] These early projects funded by the FIP do not fall in the scope of the PBS impact evaluation.

[8] Business practices adoption is defined as whether the respondent or business has adopted better business practices in the past 12 months. This includes business practices that improve, formalize or increase the efficiency of strategic planning, book-keeping, production processes, or commercial transactions.

[9] Technology adoption is defined as whether the respondent or business has adopted new production or information technologies in the past 12 months.

[10] Product diversification is defined as whether the respondent or business has produced/sold new products or crops in the past cycle.

[11] Value chain integration is defined as whether the respondent or business has experienced additional or enhanced interactions with other actors in the value chain in terms of: (1) information or (2) sales. Like other outcomes, this outcome indicator covers any interactions over the past 12 months.

Across all three value chains, the evaluation is designed to examine the impact of offering PBS for approximately one year. In addition, for the horticulture and dairy chains, the evaluation is designed to examine the differential impact of offering PBS for two years instead of one year on producers' employment creation and income, as well as on the intermediate outcomes listed above. For the handicraft chain, the evaluation is designed to examine the impact of offering PBS for two years.[12] We will discuss the specific research questions for each value chain in Section D.

C.   PDP IMPACT EVALUATION DESIGN

Given the diversity of productive sectors that PDP will target, Mathematica, MCC, and FOMILENIO have agreed that the PBS impact evaluation should be limited to three value chains: (1) handicrafts, (2) dairy, and (3) horticulture. All stakeholders agreed that these three value chains are likely to yield impacts within one to two years, with dairy being an especially important component of the Northern Zone's economy. The evaluation design is common for all three value chains, but the evaluation covers slightly different time periods according to implementation activities specific to each value chain. First, we present the core evaluation design and then we explain how this design will be adapted for each value chain.

Core Design. Our design for evaluating PBS activity is a *randomized rollout design.* This design was accepted by all the stakeholders (MCC, FOMILENIO, and Chemonics). It offers the key advantage of randomized studies: when implemented well, random assignment leads to the creation of two virtually identical groups on average at baseline, the sole difference being that only one group (the intervention group) is offered the intervention, while the other group (the control group) is not. As a result, any changes observed between the two groups following randomization can be attributed to the effects of the intervention with a known degree of statistical precision.

Under the design, all *eligible* productive units, or producers, will be offered the PBS intervention for each of the three value chains to be evaluated. However, the timing of service delivery is randomly assigned. For example, if 800 eligible producers are recruited for a specific chain, and the implementer's target for a given intervention period or cycle (for example, an annual cycle of artisan training) is to serve only 400 producers, then the implementer can be assigned a batch of 400 producers to be served in the first cycle (the intervention group) and another batch of 400 producers to be served in a subsequent cycle (the control group). Figure 1

---

[12] Impact estimation for the handicrafts chain is distinct from the other two chains because implementation plans for Phase II did not include providing services to the control group.

illustrates the random assignment of producers into these two evaluation cycles. We allow for random assignment of groups, such as artisan and dairy cooperatives and horticultural producer associations, because efficient implementation requires that Chemonics serve entire groups of producers during the same cycle.

FIGURE 1

ILLUSTRATION OF CORE RANDOMIZED ROLLOUT DESIGN
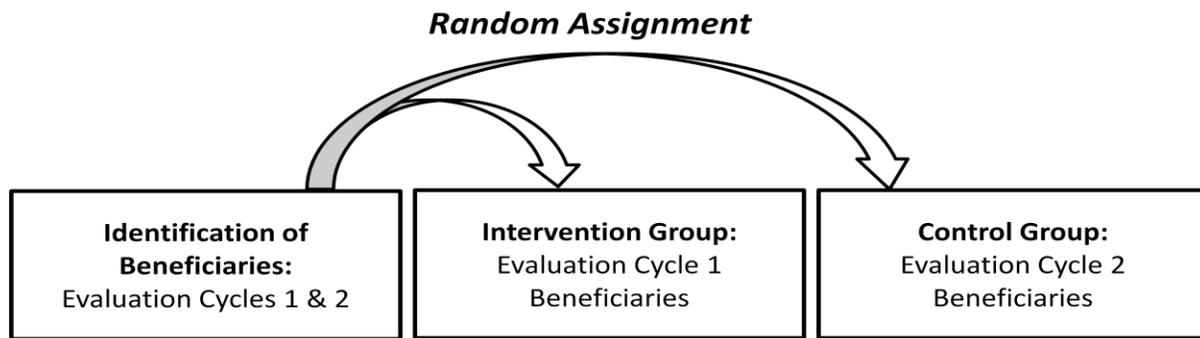
The advantage of the randomized rollout design is that all eligible producers will be offered services. The disadvantage is that impacts must be estimated before Evaluation Cycle 2 producers are offered the intervention services. The intervention cycle lengths for the three value chains to be evaluated are approximately 12 months for the handicrafts value chain, 12 months for the dairy value chain, and 10 months for horticulture.

At the end of Evaluation Cycle 1, we will estimate the impact of approximately one year of program activities for all three chains. The impact of the PBS will be defined as the difference in outcomes realized by the intervention group and the counterfactual condition (the control group). The counterfactual will not be the absence of any assistance at all, but rather the existing array of services provided in the Northern Zone, whether by the Government of El Salvador, foreign governments, financial institutions, NGOs, existing cooperatives, or other local organizations. Furthermore, at the end of Evaluation Cycle 2, we will be able to compare the effect of approximately two years of the intervention to the effect of approximately one year of the intervention for the dairy and horticulture chains. At that point, the intervention group would have had access to approximately two years of services and the control group would have had access to approximately one year of services. For the handicrafts chain, changes in implementation during Phase II allow us to compare the effect of approximately two years of the intervention (intervention group outcomes) to two years without the intervention (control group outcomes) at the end of Evaluation Cycle 2.

   a. Unit of Random Assignment. The preferred alternative is to randomly assign individual producers to be offered program services during Evaluation Cycle 1 or Cycle 2, and compare outcomes between the two groups. However, Chemonics will offer services to groups of producers rather than individual producers. For example, the handicraft intervention is currently being offered to groups of 10 to 15 producers. One of the principles on which the handicraft intervention is based is attaining some degree of association within these groups in order to become more competitive in the handicraft market. These beneficiary groups are defined in advance by the implementer, and all producers in the group will be offered services in the same implementation cycle. This arrangement requires group random assignment rather than individual assignment. Furthermore, to reduce implementation costs, Chemonics is offering the intervention services in a constrained geographic region—for example, an entire municipality—as opposed to offering the services across all municipalities. In these cases, random assignment must be conducted at the level of the municipality.

   Randomly assigning geographic clusters instead of individual producers can guard against contamination if the geographic clusters are not close to one another. There are two types of contamination. The first can occur if producers in the control group nonetheless participate in training. This could be problematic if control group members hear about training activities and find a way to attend training. The second type of contamination could occur if producers who participate in training share the techniques they learned with producers in the control group.[13]

   A disadvantage of randomly assigning groups or geographic clusters instead of individual producers is that larger samples are needed to detect impacts of the intervention. This is because producers in the same cluster—a municipality, for example—might be exposed to similar idiosyncratic influences and therefore the individual producers cannot be considered statistically independent. The relevant sample size to assess the likelihood that the study will be able to detect true impacts is therefore the number of clusters, not the number of individuals. This means that the evaluation will detect only large impacts in employment creation and household income. The section on estimating program impacts describes this issue in more detail.

   An additional consideration is that because all eligible producers within a cluster—for example, a municipality—will be in either the intervention or control group, the impact of the intervention is confounded with the municipality. We will be able to isolate the effects of the municipality (or other clustering) from the effects of the intervention to the extent that the characteristics of municipalities (such as poverty level, political affiliation, and road accessibility) vary within the intervention or control groups. Thus, the higher the number of

---

[13] Either of these types of contamination would be problematic for the evaluation because we would be unable to compare those who were offered services to those who were not offered services.

municipalities available for randomization, the better chance we have to avoid confounding the effect of the intervention with any effect attributable to municipality characteristics.

  b. Design Implementation. The implementation of the randomized rollout design consists of the following eight steps (see Figure 2).[14]

1. Chemonics **identifies or recruits potential beneficiaries**. In this first and critical step, Chemonics recruits enough producers to fill the service slots available for Evaluation Cycle 1 and enough of Evaluation Cycle 2 to populate the study sample.

2. **Lists of potential beneficiaries are available for the evaluators**. For each value chain, a single date a few weeks prior to the start of Evaluation Cycle 1 is agreed upon by stakeholders. In addition, the number of potential beneficiaries required for each value chain is agreed upon by Mathematica, MCC, FOMILENIO, and Chemonics based on Chemonics' target number of producers to be served in each implementation cycle, as well as preliminary calculations of the size of the impacts that the evaluation would be likely to detect with those sample sizes.[15] Next, Chemonics composes and shares complete lists of all eligible producers for each value chain. The lists have identifiers for each potential beneficiary and any additional information on exceptions, constraints, and relevant stratifying variables for each value chain. Exceptions might be potential beneficiaries that *must* be served in the first evaluation cycle. These producers will be excluded from the evaluation because no valid counterfactual can be identified. Constraints might be potential beneficiaries that will have to be assigned to the same evaluation cycle, such as producers in the same geographic area (for example municipality) or in the same association. In this instance, we randomly assign the entire group or geographic cluster instead of separately assigning the individuals within the group. Finally, to ensure that the Evaluation Cycle 1 and Evaluation Cycle 2 groups are balanced with regard to important characteristics—some of which are associated with outcomes— we need additional information about potential stratifying variables, such as the size of potential beneficiary groups.

3. Mathematica **randomizes the set of potential beneficiaries** into two groups: the intervention group, which will be served in Evaluation Cycle 1, and the control group,

---

[14] Here we present the steps required for all the value chains, but the schedule for each value chain varies. Each schedule was defined in collaboration with all the stakeholders involved (FOMILENIO, MCC, Chemonics, and Mathematica) in order to respect Chemonics' implementation plans as much as possible.

[15] For the handicraft value chain, the optimal sample size was 800 potential beneficiaries; for the dairy value chain it was 900 potential beneficiaries; and for the horticulture value chain it was 700 potential beneficiaries. The final sample sizes were smaller however, due to large proportions of exceptions in each chain.

which will be served in Evaluation Cycle 2. The randomization procedure takes into account the exceptions and constraints discussed above. The evaluator transmits the list of assigned potential beneficiaries to Chemonics on the agreed dates for each value chain in order for Chemonics to communicate to the producer whether they would be served right away or during the next evaluation cycle.

4. **Collect baseline data** from all eligible producers before the start of Evaluation Cycle 1. The Dirección General de Estadística y Censos in El Salvador (DIGESTYC) collects baseline data from all potential beneficiaries before each intervention starts. Although data collection could extend up to two months beyond the start of Evaluation Cycle 1 service activities, ideally all baseline data should be collected before the start of Evaluation Cycle 1 service delivery. Dates and more specific information are presented in the data collection section and summarized in Table 3.

5. During Evaluation Cycle 1, Chemonics **offers the intervention services to the intervention group, but not to the control group**. Mathematica communicates with Chemonics and FOMILENIO to monitor the implementation of the intervention and identify potential problems—such as contamination, among others—in order to deal with these problems in a timely manner.

6. **Collect first follow-up data** close to the end of Evaluation Cycle 1. The specific dates for the first follow-up survey vary by chain as presented in Table 3 in the data collection section. The dates were selected to be as late as possible within Evaluation Cycle 1, keeping in mind that potential beneficiaries to whom the intervention will be offered in Evaluation Cycle 2 are waiting to receive services. Thus, the duration of the field period for the first follow-up survey is constrained by the duration of the interval between the evaluation cycles for each value chain (see Figure 2).

7. **Monitor the implementation of the intervention in Evaluation Cycle 2**. MCC and FOMILENIO expressed interest in assessing the impact of the intervention at the end of Evaluation Cycle 2. In this case, the evaluation will provide a comparison between receiving two years of the intervention services (group assigned to Evaluation Cycle 1) and receiving one year of the intervention services (group assigned to Evaluation Cycle 2).[16] Mathematica and Chemonics will monitor the implementation during Evaluation Cycle 2 in order to identify potential problems and address them in a timely manner.

---

[16] The exception is the handicrafts chain, for which the evaluation will provide a comparison between receiving two years of services and receiving no services.

8. **Collect second follow-up data** close to the end of Evaluation Cycle 2. The specific dates for the second follow-up survey vary by value chain as presented in Table 3 in the data collection section.

FIGURE 2

SEQUENCE OF ACTIVITIES OF THE CORE RANDOMIZED ROLLOUT DESIGN
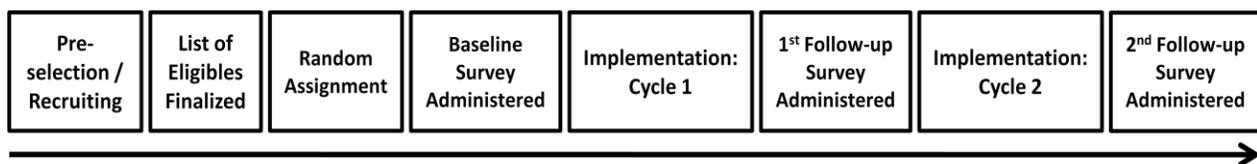
By 2010, the first five activities were completed for each of the value chains in the evaluation:

- **Dairy chain.** Mathematica randomized 28 groups with 595 eligible individuals into treatment and control groups: 14 productive groups with 295 eligible individuals were randomly selected to be offered Chemonics services beginning in May 2010 (treatment group), and 14 productive groups with 300 eligible individuals were randomly selected to be offered Chemonics services beginning in May 2011 (control group).[17] In May 2010, DIGESTYC finished baseline data collection for the dairy value chain.

- **Horticulture chain.** Mathematica randomized 31 groups with 647 eligible individuals into treatment and control groups: 15 productive groups with 324 eligible individuals were randomly selected to be offered Chemonics services beginning in April 2010 (treatment group), and 16 productive groups with 323 eligible individuals were randomly selected to be offered Chemonics services beginning in March 2011 (control group). [18] In June 2010, DIGESTYC finished baseline data collection for the horticulture value chain.

- **Handicraft chain.** Mathematica randomized 19 municipalities with 674 eligible individuals into treatment and control groups: 9 municipalities with 337 eligible individuals were randomly selected to be offered Chemonics services beginning in September of 2009

---

[17] Twenty individual producers were organized into two productive groups for the purpose of randomization. As a result, 11 individuals were randomized jointly into control, and 9 individuals were randomized jointly into treatment.

[18] The original list of eligible beneficiaries was 945 eligible individuals in 48 groups: 298 eligible individuals in 17 groups were classified as exceptions, and therefore not randomized.

(treatment group), and 10 municipalities with 337 eligible individuals in were randomly selected to be offered Chemonics services beginning in September of 2011 (control group).[19] DIGESTYC collected baseline data for the handicraft value chain in late 2009. In November 2010, DIGESTYC completed follow-up data collection for the handicraft value chain following approximately one year of services (for the treatment group only).

D.   OUTCOME INDICATORS AND DATA SOURCES

The PBS impact evaluation will assess both main and secondary outcomes resulting from the offer of intervention activities. These outcomes are described in more detail below.

1.   Impact Evaluation Outcomes

a. Main Outcomes. The ultimate goal of PBS is to increase the employment and household income of producers in El Salvador's Northern Zone. In particular, we will collect information on sources of income that are most directly affected by the training programs, specifically income from handicraft, dairy, and horticulture production. We will also track employment information, measured by the number of people paid by producers in the study sample to perform productive activities. Table 1 summarizes the evaluation's two main outcomes and their corresponding indicators.[20] We will collect data on these outcomes during the baseline, first follow-up, and second follow-up surveys.

b. Secondary Outcomes. In addition to employment and income outcomes, we will closely examine secondary outcomes through which the training programs are intended to improve household income, including production levels, sale prices, adoption of new practices and technologies, as well as enhanced product diversification and value chain integration. Table 1 summarizes the evaluation's key secondary outcomes and their corresponding indicators. As with the study's main outcomes, we will collect data on these secondary outcomes during the baseline, first follow-up, and second follow-up surveys.

---

[19] The original list of eligible beneficiaries included 750 eligible individuals in 22 municipalities. Three municipalities with 76 eligible beneficiaries were exceptions and were not included in the randomization.

[20] An outcome indicator is a metric that quantifies an outcome of interest in a specified time frame. In the case of productive development indicators, this time frame is one evaluation cycle. For the handicraft and dairy value chains, one evaluation cycle is 12 months long. For the horticulture value chain, one evaluation cycle is 10 months.

TABLE 1

KEY PBS MAIN AND SECONDARY OUTCOMES

Outcome	Indicator	Time of Collection
Main		
Employment	Number of full-time equivalent jobs (250 *jornadas*) created	Baseline, first, and second follow-ups[21]
Income	Net household income in past cycle, including income from productive development	Baseline, first, and second follow-ups
Secondary		
Production levels	Amount of production (bottles of milk, kilograms of fruits/vegetables, number of handicrafts, etc.)	Baseline, first, and second follow-ups
Sale Prices	Price at which production was sold	Baseline, first, and second follow-ups
Business Practices	Group or individual beneficiary has composed a business plan, used basic bookkeeping practices, taken quality control measures and/or taken an inventory	Baseline, first, and second follow-ups
Technology Adoption	Group or individual beneficiary has used new (information) technologies to produce/sell products in past cycle	Baseline, first, and second follow-ups
Diversification	Number of product types (or crops) produced; amount of crops/products produced/sold	Baseline, first, and second follow-ups
Value Chain Integration	Number of different sources of information regarding prices and client preferences; number of different buyers in the past cycle	Baseline, first, and second follow-ups

PDP = Productive Development Project.

2.   Encuestas de Desarrollo Productivo

Most of the data needed to construct these key outcome indicators cannot be collected using national surveys or administrative records. To estimate the impact of PBS on employment creation and income, we must collect primary data on baseline characteristics, outcomes, and utilization of PBS for producers in the study sample. We will collect these primary data through the Encuestas de Desarrollo Productivo (EDPs), which are surveys developed specifically for this impact evaluation. The EDPs will be administered to all eligible producers in the three value chains featured in the evaluation; this includes all producers in the intervention group as well as the control group.

---

[21] The baseline handicraft survey provides only partial information on this outcome, but all follow-up instruments will feature full information on employment creation.

MEMO TO: Rebecca Tunstall and Claudia Argueta
FROM: Larissa Campuzano and Randall Blair
DATE: 2/16/2011
PAGE: 13

Mathematica has developed EDP instruments for each of the three value chains included in the evaluation. FOMILENIO has contracted DIGESTYC to field the baseline EDP for each value chain. In future years, DIGESTYC is expected to administer two follow-up rounds of the EDP for each value chain.

a. Survey Instruments. Both baseline and follow-up versions of the EDP will capture data on the key outcomes mentioned above; baseline surveys will measure producers' initial practices, production, employment, and income, whereas follow-up surveys will monitor how producers' practices, production, employment, and income change throughout the study period. In addition, the EDP will also collect background and participation data. Background data include demographic information about individuals and their communities. Participation data will detail producers' participation in PDP services, as well as technical and financial assistance from sources other than PDP. Combined with outcome data, these background and participation data will provide a comprehensive picture of producers' characteristics, resources, and production over the course of the study.

Within each productive chain, DIGESTYC will administer two distinct survey instruments: an individual instrument and a group leader instrument.[22] All producers on master lists will complete the individual instrument, but only leaders of cooperatives, workshops, and other productive groups will complete the group leader instrument. The focus of the individual instrument is each respondent's productive activities as well as his or her household costs and income. The focus of the group instrument is the group's collective activities, costs, and income. Table 2 describes the topics covered in the individual instrument versus the group leader instrument.

---

[22] Mathematica developed an individual and leader survey for each of the study's three value chains, for a total of six unique survey instruments.

MEMO TO: Rebecca Tunstall and Claudia Argueta
FROM: Larissa Campuzano and Randall Blair
DATE: 2/16/2011
PAGE: 14

TABLE 2

TOPICS COVERED BY THE EDP

Individual Instrument	Leader Instrument
• Demographic information	• General group information
• Household composition	• Group production and sales
• Productive activities	• Information and marketing
• Productive expenses/income	• Common problems facing the group
• Productive activities (Horticulture and Dairy)[a]	• Group productive activities (Handicrafts)
• Household income	• Group expenses/income
• Household expenses	
• Credit	

[a] To capture information on the primary level of production, productive activities were measured at the individual level for horticulture and dairy, and the group level for handicrafts.

b. Survey Sample Frames and Sampling Plan. The target populations for each of the three value chains are eligible producers as determined by Chemonics. In the handicrafts and horticulture chains, these eligible producers are organized in productive groups. In the dairy chains, some eligible producers are members of productive groups, but others are not. In all chains, Chemonics identified producers through a formal recruitment process for each value chain. First, Chemonics held a series of meetings with artisans, dairy producers, and farmers in various municipalities in the Northern Zone. Second, Chemonics staff compiled a master list of all interested *and* eligible producers for each value chain. These master lists comprised the complete sample frames for the EDP. Each eligible beneficiary included in these master lists will be asked to complete a baseline interview *and* two follow-up interviews throughout the course of the study. In addition, a group leader from each productive group included in the master lists will be asked to complete a baseline interview and two follow-up interviews throughout the course of the study. Because the same individuals and groups will be interviewed up to three times over the course of the evaluation, the EDP will yield a longitudinal data set of PDP eligible producers and productive groups, although there might be cases in which producers drop out from the intervention and cannot be located or refuse to respond to the follow-up surveys.

The municipalities represented in master lists vary based on demand for each value chain's services. Nineteen municipalities were represented in the master list of eligible producers for the handicraft value chain; 17 municipalities were represented in the master list for the horticulture chain; and 22 municipalities were represented in the master list for the dairy value chain.

c. Data Collection Plan. Under its agreement with FOMILENIO and DIGESTYC, Mathematica is responsible for drafting all EDP survey instruments and manuals, training all data collectors (or providing proper oversight of training), cleaning all three list frames, and randomizing potential beneficiaries for each value chain. For each value chain, Mathematica staff will provide DIGESTYC with cleaned sample frames. DIGESTYC will administer all baseline and follow-up surveys according to these sample frames.

DIGESTYC will administer baseline and follow-up surveys for each of the study's three value chains. For each baseline and follow-up EDP, DIGESTYC will:

- Revise and diagram the survey instruments and administer field tests
- Select surveyors to administer the surveys according to established criteria
- Provide a locale and equipment for training
- Provide surveyors with global positioning system training
- Prepare all survey maps and materials, excluding training manuals
- Provide all information required for data quality review
- Administer the survey in the field according to the cleaned sample frame
- Review and code data, and provide quality control
- Compose, verify, and submit a database of survey data
- Submit monthly progress reports and a final report

DIGESTYC estimates that the following the personnel are required to conduct each survey:

- 13 administrative staff, including a coordinator and field supervisor
- 15 staff for surveyor teams, including 3 supervisors, 9 surveyors, and 3 drivers
- 6 staff for information processing

DIGESTYC surveyor teams—each comprising one supervisor and three surveyors—will travel to the location of producers' cooperatives, workshops and groups to survey all producers in the sample frames.[23] Working with Chemonics, the surveyor teams will notify each group of the date and time they will hold a meeting to interview all group members. The majority of producers will be interviewed during these meetings. Following the meetings, DIGESTYC surveyors will contact and/or travel to the homes of producers that did not attend the meetings in an effort to interview all individuals in the sample frame. Surveyors may also travel to the homes of producers to speak with other members of the producers' households that are better informed about specific household costs and income, such as agricultural income. In an effort to control survey costs, DIGESTYC will track and attempt to minimize all transportation costs associated with locating and interviewing individuals outside of group meetings.

---

[23] If such a location does not exist, interviews will take place in a community building or a private home.

The target response rate for baseline surveys for each value chain is 88 percent, a figure proposed by DIGESTYC given its experience with other baseline surveys it has conducted for FOMILENIO. The response rate for follow-up surveys may be lower, as locating and interviewing eligible producers will become more difficult as the study progresses. DIGESTYC will provide regular updates of survey response rates during the survey's field phase.

As described above, the timing of data collection largely depends on the start- and end-dates of the evaluation cycles of each of the value chains. Because the handicraft value chain cycle began prior to the dairy and horticulture chains, DIGESTYC administered the handicraft baseline survey in October 2009. The dairy survey was administered in May 2010, and the horticulture baseline survey was administered in June 2010. Follow-up surveys for the handicraft and dairy chains will be administered on a 12-month cycle; follow-up surveys for the horticulture chains will be administered on a 10-month cycle. Table 3 outlines additional key dates related to PDP data collection, including dates of follow-up surveys.

TABLE 3

KEY DATA COLLECTION AND EVALUATION DATES, BY VALUE CHAIN

		Value Chain		
		Handicrafts	Horticulture	Dairy
Baseline Survey				
1	Revised baseline instrument to DIGESTYC	August 2009	February 2010	February 2010
2	List of potential beneficiaries sent by Chemonics to Mathematica	September 2009	March 2010	March 2010
3	Select treatment and control groups	September 2009	March 2010	March 2010
4	Conduct interviewer training	September 2009	April 2010	April 2010
5	Baseline data collection	October 2009	June 2010	May 2010
6	Baseline data set sent by DIGESTYC to Mathematica	April 2010	September 2010	September 2010
7	Data review	April 2010	September 2010	September 2010
8	Baseline data collection documentation sent by DIGESTYC to Mathematica	April 2010	September 2010	September 2010
First Follow-up Survey				
9	Develop draft follow-up survey instrument	August 2010	January 2011	March 2011
10	Revise instrument based on comments from Chemonics and FOMILENIO	September 2010	February 2011	April 2011
11	Conduct interviewer training	October 2010	May 2011	June 2011
12	Follow-up (Round 1) data collection	November 2010	May 2011	June 2011
Second Follow-up Survey				
13	Develop draft follow-up survey instrument	June 2011	February 2012	March 2012
14	Revise instrument based on comments from Chemonics and FOMILENIO	July 2011	March 2012	April 2012

	Value Chain		
	Handicrafts	Horticulture	Dairy
15 Conduct interviewer training	August 2011	April 2012	May 2012
16 Follow-up (Round 2) data collection starts	September 2011	May 2012	June 2012

E. ESTIMATING PROGRAM IMPACTS

Random assignment ensures that, on average, producers in the intervention group and producers in the control group are balanced on all characteristics before the beginning of the intervention. Hence, after Evaluation Cycle 1, the difference between the mean of the outcome of interest for the intervention group and the mean of that same outcome for the control group yields an unbiased estimate of the impact of the offer of PBS. The precision of the impact estimates depends mainly on the sample sizes allocated to the treatment and control groups; however, this precision can be improved by controlling for other covariates in a regression model. Regression adjustment can also help account for any differences between the treatment and control groups in baseline characteristics that arose by chance.

1. Impact Estimation

a. Core Regression Specification for Each Value Chain. The impact analysis will rely on a core regression specification for each value chain. In this specification, we have assumed that we will randomize groups or clusters of producers in each value chain. The econometric specification compares how groups or clusters in the treatment group changed over time with how groups or clusters in the control group changed over time, controlling for idiosyncratic differences in the two groups. The basic model can be expressed as follows:

$$(1) \quad y_{igt} = \alpha + \beta' x_{igt-1} + \delta y_{igt-1} + \lambda T_g + \eta_g + \varepsilon_{igt}$$

where y_{igt} is the outcome of interest for beneficiary i in group or cluster g at time t; x_{igt-1} is a vector of baseline characteristics of beneficiary i in group or cluster g (note that these characteristics could be both time-invariant, such as gender, or time-variant, such as time worked); y_{igt-1} is the baseline value of the outcome indicator of beneficiary i in group or cluster g; T_g is an indicator equal to one if group or cluster g is in the treatment group and zero if it is in the control group; η_g is a group-specific term that could be modeled as a group or cluster "random effect" or as a fixed effect; and ε_{igt} is a random error term for beneficiary i in group or cluster g observed at time t. The parameter estimate for λ is the estimated impact of the program for each value chain.

The vector of baseline characteristics x_{igt-1} will include both beneficiary and group level characteristics. We will control for group characteristics, such as size of the group, average income at the group level, etc. We will also control for producer characteristics, such as level of education, gender, age, number of years working in the productive chain, and so on.

b. Pooling Impact Estimates across Value Chains. To provide a measure of the overall impact of PDP, we can pool the impacts of the three value chains. This can be done by aggregating the estimates calculated by the models specified in Equation (1) for each of the value chains into a weighted average (similar to what is done when site impacts are pooled into one overall impact estimate). Alternatively, we could specify a similar model to Equation (1) that would use the data for the three value chains and would obtain one pooled impact estimate. However, obtaining a pooled impact estimate presents some challenges. The interventions across value chains are not homogeneous; each intervention was designed to address the needs of that particular value chain and was implemented differently. Although income is the primary outcome measure for the three value chains, the offered services are inherently different, as are the productive activities these services support. Furthermore, the interventions have a different implementation schedule across the value chains; this further reduces the intervention's homogeneity across chains. Most importantly, for the results after Evaluation Cycle 2, the analysis for the dairy and horticulture value chains will compare the offer of PBS for two years versus the offer of PBS for one year, while the handicrafts value chain will compare the offer of PBS services for two years versus no offer of PBS. Therefore, our recommendation is to focus on impacts for each individual value chain. However, as a sensitivity analysis, we will consider pooling either the chain-specific impact estimates or the data across value chains to produce a single impact estimate.

c. Intent-to-Treat and Treatment-on-the-Treated Estimates. For each of the three value chains in the evaluation, we will produce two types of impact estimates: intent-to-treat estimates and treatment-on-the-treated estimates. Intent-to-treat estimates capture the impact of the offer of PBS, regardless of whether the producers in the treatment group accepted this offer. To construct these estimates, we will compare the entire sample of producers in treatment to the entire sample of producers in control. Because the stakeholders are also interested in the effect of the treatment on those who participated in the intervention, we will estimate the treatment-on-the-treated effects. To construct these estimates, we will use an instrumental variable approach in which the random assignment results will serve as the instrument. This variable is a good candidate for an instrument because being assigned to the intervention or control group will be highly correlated with receiving PBS. In addition, because the assignment was random, it will be uncorrelated with the outcomes.

MEMO TO: Rebecca Tunstall and Claudia Argueta
FROM: Larissa Campuzano and Randall Blair
DATE: 2/16/2011
PAGE: 19

2.   Statistical Power

For our power calculations, we made two key assumptions: (1) a total attrition rate of 15 percent for both groups and individuals at the end of the three-year follow-up period; (2) and the percentage of variance explained by the regression model (R^2) equal to 0.7 for both groups and individuals.[24] For all three value chains, we applied the intra-class correlations for variables from the three baseline EDPs (handicrafts, horticulture, and dairy). In addition, we based our calculations on the means of constructed variables from the three EDPs, as well as their standard deviations.

We calculated the statistical power that we will be likely to have after three years of follow-up, given the study's current sample sizes, the number of randomized clusters, and our assumptions above. The target sample sizes agreed upon by all the stakeholders were 800 potential eligible producers for the handicraft chain, 900 for the horticulture chain, and 700 for the dairy chain. After exceptions and ineligible producers were excluded, the actual sample sizes were 674 artisans in 19 municipalities; 647 horticulture producers in 31 groups; and 595 dairy producers in 28 groups. Given these sample sizes—and more importantly, the number of groups or municipalities randomized in each sample—we calculated the estimated size of the effects we are likely to detect at follow-up for the intent-to-treat analysis. These are: (1) the minimum detectable effect size (MDE), which is measured in standard deviations of the outcome of interest, household income, relative to the control group; and (2) the equivalent minimum detectable impact (MDI), which is measured as the percentage increase in household income relative to the control group. The MDE and the MDI are the minimum change in the standard deviation and percentage of household income, respectively, that we would be able to detect given parameters discussed above.

For the handicraft chain, we randomly assigned 674 producers in 19 municipalities into 9 municipalities in the treatment group and 10 municipalities in the control group. Under this arrangement, we will be able to detect an income change of 0.27 of a standard deviation (MDE), which translates to a 33 percent change in income (MDI; see Table 4). For the horticulture value chain, we randomly assigned 647 producers in 31 groups into 15 groups in treatment and 16 groups in control. Under this arrangement, we will be able to detect an income change of 0.18 of a standard deviation (MDE), which translates to a 26 percent change in income (MDI). For the dairy value chain, we randomly assigned 595 producers in 28 groups into 14 groups in treatment and 14 groups in control. Under this arrangement, we will be able to detect an income change of 0.17 of a standard deviation (MDE), which translates to a 22 percent change in income (MDI). Pooling all three value chains—in which 78 groups and 1,916 producers are randomized into

---

[24] The R^2 of 0.7 was corroborated by our analysis of baseline handicrafts data. This high R^2 was achieved by including municipal-level stratifying variables in regressions. We are also assuming 80 percent power, 5 percent significance, and a two-tail test.

treatment and control—we calculated an MDE of 0.12 and an MDI of 17 percent. In Appendix A, we discuss these power calculations in greater depth, including recent changes in estimates and assumptions.

TABLE 4

ESTIMATED DETECTABLE EFFECTS AND IMPACTS ON NET HOUSEHOLD INCOME,
BY VALUE CHAIN

Random Assignment	MDE (Fraction of a standard deviation)	MDI (Percentage Change in Income)
Handicrafts		
19 municipalities (674 producers)	0.27	33
Horticulture		
31 groups (647 producers)	0.18	26
Dairy		
28 groups (595 producers)	0.17	22
Combined Sample		
78 groups (1,916 producers)	0.12	17

Source: Mathematica calculations based on data from the 2009-2010 Encuestas de Desarrollo Productivo – (EDP)

MDE = minimum detectable effect; MDI = minimum detectable impact.

3.   Specific Impact Questions

In this section, we list the specific research questions we can answer regarding the impact of PBS for each of the three value chains, with an emphasis on distinguishing Phase I and Phase II services. For the handicrafts value chain, the impact of Phase I PBS can be isolated from the impact of Phase II PBS because Phase II services started as the first follow-up survey was completed in November 2010. For the horticulture and dairy chains, the impact of Phase I services is not easily disentangled from the impact of Phase II services because follow-up surveys will capture a mix of Phase I and Phase II services. To be fielded in May 2011, the first follow-up survey for the horticulture chain will allow us to measure the impact of all services provided under Phase I, in addition to the first six months of Phase II services (provided that Phase II services began in early November 2010). To be fielded in June 2011, the first follow-up survey for the dairy chain will allow us to measure the impact of all services provided under Phase I, in addition to the first six months of Phase II services (provided that Phase II services began in early November 2010).

The evaluation will answer the following questions for the handicraft value chain:

1) What was the effect of 12 months of services (from mid-September 2009 to mid-September 2010) versus no services? We will calculate this by comparing treatment and control outcomes from the first follow-up survey.

2) What was the effect of 24 months of services (from mid-September 2009 to mid-September 2011) versus no services? We will calculate this by comparing treatment and control outcomes from the second follow-up survey.

We can answer the following questions for the dairy value chain:

1) What was the effect of six months of PBS services under Phase I (from May to October 2010) versus no services? We will calculate this by comparing treatment and control outcomes for the 2010 rainy season (May to October 2010). Key indicators for this analysis are productive income, quantities sold, and price of sold products.

2) What was the effect of six months of PBS services under Phase II (from November 2010 to April 2011) following six months of PBS services under Phase I versus no services? We will calculate this by comparing treatment and control outcomes for the 20010-2011 dry season (November 20010 to April 2011). Key indicators for this analysis are productive income, quantities sold, and price of sold products.

3) What was the overall effect of 12 months of PBS services, regardless of the Phase, (from May 2010 to April 2011) versus no services? We will calculate this by comparing treatment and control outcomes using data from the first follow-up survey.

4) What was the effect of 24 months of PBS services (from May 2010 to April 2012) versus 12 months of PBS services (from May 2011 to April 2012)? We will calculate this by comparing treatment and control outcomes using data from second follow-up survey.

We can answer the following questions for the horticulture value chain:

1) What was the effect of seven months of Phase I services (from April to October 2010) versus no services? We will calculate this by comparing treatment and control outcomes for the 2010 rainy season (May to October 2010). Key indicators for this analysis are productive income, quantities sold, and price of sold products.

2) What was the effect of four months of Phase II services (from November 2010 to February 2011) following six months of Phase I services versus no services? We will

calculate this by comparing treatment and control outcomes for the 2010-2011 dry season (November 2010 to April 2011). Key indicators for this analysis are productive income, quantities sold, and price of sold products.

3) What was the overall effect of 11 months of PBS services (from April 2010 to February 2011), regardless of the phase, versus no services? We will calculate this by comparing treatment and control outcomes using data from the baseline and first follow-up survey.

4) What was the effect of 24 months of services (from April 2010 to March 2012) versus 13 months of services (from March 2011 to March 2012)? We will calculate this by comparing treatment and control outcomes using data from second follow-up survey.

F.   REPORTING PLANS

Over the course of the evaluation, we will provide three reports summarizing the findings across all three value chains. The baseline report will summarize findings from the baseline EDPs and will analyze the characteristics of the intervention group versus the control group in all three value chains. The first follow-up report will summarize the findings from the first round of follow-up surveys, which will be administered after the intervention group has received one cycle of services and the control group has received no services. The main focus of this report is to quantify the impact of one cycle of productive development services on producers' incomes, employment and other outcomes. The second follow-up report will cover the findings from the second follow-up surveys, which will be administered after the intervention group has received two cycles of services and the control group has received one cycle of services.[25] The main focus of this final report is to quantify the impact of two cycles of productive development services versus one cycle of services. Table 5 presents tentative dates of all key deliverables associated with the PDP impact evaluation.

---

[25] The exception to this is the handicrafts chain, in which the intervention group will have received two cycles of services and the control group will have received no services.

TABLE 5

KEY DELIVERABLES AND DATES OF SUBMISSION

Deliverable	Date
Baseline Report	March 2011
Interim Report (First Follow-Up)	December 2011
Final Report (Second Follow-Up)	January 2013

cc: Claudia Argueta (FOMILENIO), Ricardo Orellana (FOMILENIO), Francisco Munguía (DIGESTYC), Sarah Crawford (MCC), Kenny Miller (MCC), Carlos Domínguez (Chemonics), William Mejía (Chemonics), File

Appendix A

This appendix provides a brief discussion of the statistical power for the impact evaluation of the Productive Development Project (PDP). For this evaluation, statistical power is our ability to detect increases in income and other outcomes with high probability, given that these increases are real. We analyze statistical power for each of the three value chains in the evaluation: handicrafts, dairy, and horticulture.

For our power calculations, we made two key assumptions: (1) a total attrition rate of 15 percent for both groups and individuals at the end of the three-year follow-up period; (2) and the percentage of variance explained by the regression model (R^2) equal to 0.7 for both groups and individuals.[26] For all three value chains, we applied the intra-class correlations for variables from the three baseline EDPs (handicrafts, horticulture, and dairy). In addition, we based our calculations on the means of constructed variables from the three EDPs, as well as their standard deviations.

We originally calculated the statistical power that we will be likely to have after three years of follow-up, given the study's current sample sizes, the number of randomized clusters, and our assumptions above. The target sample sizes agreed upon by all the stakeholders were 800 potential eligible producers for the handicraft chain, 700 for the dairy chain, and 900 for the horticulture chain. After exceptions[27] and ineligible producers were excluded, the actual sample sizes were 674 artisans in 19 municipalities; 595 dairy producers in 28 groups; and 647 horticulture producers in 31 groups. Given these sample sizes—and more importantly, the number of groups or municipalities randomized in each sample—we calculated the estimated size of the effects we are likely to detect at follow-up. These are: (1) the minimum detectable effect size (MDE), which is measured in standard deviations of the outcome of interest, relative to the control group; and (2) the equivalent minimum detectable impact (MDI), which is measured as the percentage increase in the outcome of interest among the treatment group relative to the control group. Examining the outcome of household income, the MDE and the MDI are the minimum change in the standard deviation and percentage of household income, respectively, that we would be able to detect given parameters discussed above.

We calculated that the minimum detectable change (MDE) in net household income that we will be able to measure with high probability is 0.27, 0.18, and 0.17 standard deviations of household income for the handicraft, horticulture, and dairy value chains, respectively. This is

---

[26] The R^2 of 0.7 was corroborated by our analysis of baseline handicrafts data. This high R^2 was achieved by including municipal-level stratifying variables in regressions. We are also assuming 80 percent power, 5 percent significance, and a two-tail test.

[27] In memo ESVED-206, exceptions are defined as "potential beneficiaries that *must* be served in the first evaluation cycle" due to key characteristics or logistical issues.

equivalent to changes in household income (MDI) of 33, 26, and 22 percent for the handicraft, horticulture, and dairy chains. Combining all three value chains, we calculated an MDE of 0.12 and an MDI of 17 percent. Table A.1 provides MDIs for a variety of outcome variables that are relevant to the PBS impact evaluation, including poverty levels and the number of full-time equivalent jobs (FTEs) generated by producers in the sample.

TABLE A.1

ESTIMATED MINIMUM DETECTABLE EFFECTS (MDEs) AND MINIMUM DETECTABLE IMPACTS (MDIs) ON KEY ANNUAL INDICATORS, BY VALUE CHAIN

Level	Indicator	Handicrafts		Horticulture		Dairy		Combined	
		MDE	MDI	MDE	MDI	MDE	MDI	MDE	MDI
Household	Gross income	0.30	37%	0.21	28%	0.21	26%	0.14	20%
	Net income	0.27	33%	0.18	26%	0.17	22%	0.12	17%
	Gross productive income	0.47	76%	0.25	35%	0.24	27%	0.19	37%
	Net productive income	0.46	71%	0.32	52%	0.19	29%	0.19	39%
	Consumption	0.32	24%	0.18	13%	0.19	15%	0.13	11%
	Under $1.25 poverty line (percentage based on consumption)	0.30	14%	0.17	5%	0.13	2%	0.13	4%
	Under $1.25 poverty line (percentage based on income)	0.24	10%	0.17	7%	0.18	4%	0.11	4%
Individual	Gross productive income	0.46	74%	0.26	36%	0.23	25%	0.19	37%
	Net productive income	0.44	65%	0.29	52%	0.18	25%	0.19	39%
	Total gross income	0.33	48%	0.21	32%	0.21	28%	0.15	26%
	Total net income	0.31	44%	0.19	30%	0.21	29%	0.14	24%
	Percentage of producers that employ workers	0.31	9%	0.18	9%	0.18	9%	0.13	6%
	Number of FTEs generated	0.21	NA[a]	0.12	0.15	0.19	0.46	0.10	0.20
	Number of people paid for production	0.30	0.44	0.12	0.31	0.19	0.35	0.09	0.21

Source: Mathematica calculations based on data from the 2009-2010 Encuests de Desarrollo Productivo

Note: For poverty measures and percentage of producers that employ workers, the impacts reported are measured in percentage points (for example, a decrease in the poverty rate of 4 percentage points). For the number of FTEs generated and people paid, the impacts reported are measured as number of FTEs and people paid, respectively (for example, an increase of 0.2 FTE jobs per year). For all other variables, the impacts reported represent percentage point changes (for example, an increase in net household income of 17 percent of the original household income).

[a] Information for the baseline period is unavailable for the handicrafts baseline survey, but will be available following completion of the first follow-up survey.

Several factors contribute to the low power—or high MDEs and MDIs—of the impact evaluation. The most important of these factors are:

1. ***High variability in household income.*** Using data from the EDPs, we found extremely high variability in producers' reported net household income. Mean net annual household income is $3,011, $3,535, and $9,615 in the handicraft, horticulture, and dairy chains, respectively, with a standard deviation (SD) of $6,857, $8,359, and $17,845, respectively. This SD-to-mean proportion (or coefficient of variation) of over 1.75 is very high, even higher than the 1.4 SD-to-mean proportion we found in data from the 2007 Encuesta de Hogares de Propósitos Múltiples (EHPM) for families in the Northern Zone. This high variability is the main cause for the study's high MDIs. The study's minimal detectable effect sizes (MDEs) of around 0.2 for key outcomes are moderate for impact studies of this nature. However, when multiplied by the very large standard deviation of annual household income, these MDEs produce very large minimum detectable impacts (MDIs) of between 26 and 33 percent of net household income.

2. ***Randomization at the group level and smaller sample sizes than anticipated.*** When doing random assignment at the group or municipal level instead of the individual level, the number of groups or municipalities randomized is a larger determinant of the MDIs than the number of individual producers in the sample. The study's current range of between 19 and 31 groups or municipalities in each value chain (with a high average number of individuals in each group) results in much lower ability to detect changes than an originally assumed range of between 40 and 50 groups or municipalities in each chain (with a lower average number of individuals in each group).[28] In addition, the sample of individuals is smaller than we first expected due to the exclusion of exceptions and ineligible producers. These sample sizes were further reduced to accommodate an impact analysis at the household level, as some individuals in the sample belong to the same household. Using data from the baseline Encuestas de Desarrollo Productivo (EDP), we found that these samples of individual producers represent only 604, 538, and 583 households in the handicraft, dairy, and horticulture chains, respectively. Therefore our ability to detect smaller changes in outcomes was further diminished due the factors described above.

3. ***Moderate intra-class correlations.*** The intra-class correlations calculated from the baseline EDPs used for these updated calculations (0.12 for gross income and 0.08 for net income for the combined sample) are higher than the intra-class correlation we used in previous estimates: the correlation of gross household income at the municipal level

---

[28] See memo ESVED-206, dated September 14, 2009, for MDE and MDI estimates in which we assumed that between 40 and 50 groups in each chain would be randomized into treatment and control

(0.03) obtained from the 2007 Encuesta de Hogares de Propósitos Múltiples (EHPM) for the Northern Zone. However, our updated intra-class correlations—for the combined sample as well as for each chain—are more appropriate than our previous estimate from EHPM data, given that they better represent the population of producers in the sample, rather than a sample of all households in the Northern Zone. These updated intra-class correlations slightly decrease the study's ability to detect changes relative to previous estimates.

The current participation rate among eligible producers in the three value chains underscores another problematic issue related to the study's statistical power. Only 53 percent of eligible handicraft producers randomized in 2009 received services from Chemonics during 2010, and similar participation rates were reported for the other two value chains. Under the study's intent-to-treat analytical approach, the producers that actually participated in productive development services—roughly half of the sample at the time of this writing—would have to experience increases in household income of much higher than 50 percent in order to detect an income increase of 50 percent across the entire study sample. This might be impossible, as these estimates imply that participants' income actually doubles over a short period (about three years following randomization). Regardless of the intervention, income increases of this magnitude are not likely, particularly under the current economic circumstances in El Salvador.