MATHEMATICA
Policy Research, Inc.

MEMORANDUM

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

TO: Rebecca Tunstall and Orlando Martínez

FROM: Larissa Campuzano, Lorenzo Moreno, and Randall Blair

DATE: 9/14/2009
ESVED-206

SUBJECT: Final Impact Evaluation Design for the Productive Development Project

---

This memorandum describes the final impact evaluation design for the Productive Development project, as agreed upon by Mathematica Policy Research, Millennium Challenge Corporation (MCC), Fondo del Milenio (FOMILENIO), and Chemonics. The document builds on our previous proposals (see memoranda ESVED-105, ESVED-134, and ESVED-170); our detailed review of Chemonics' work plan and other related documentation; Rebecca Tunstall's internal memorandum to MCC staff dated March 3, 2009; extensive discussions with you and stakeholders in El Salvador during the past months; and the agreement signed by Chemonics during our visit to El Salvador in May 2009.

EXECUTIVE SUMMARY

The final, core evaluation design is a *randomized rollout design*. Mathematica® is adapting this design to each of the three value chains that the impact evaluation will consider: (1) Artisan, (2) Dairy, and (3) Horticulture. To capture information on productive activities and the evaluation's key outcome of household income, Mathematica has designed the Encuestas de Desarrollo Productivo (EDP) which are tailored to the needs of each of the three productive chains. Baseline data collection will take place during the period October 2009-May 2010. Follow-up data collection will start in September 2010. The proposed design is flexible enough to accommodate Chemonics' implementation of the intervention's activities. If successfully implemented, the impact evaluation will detect changes of between 35 and 43 percent in household income. We will report findings for each value chain at the end of the first and second follow-up periods.

A. DESCRIPTION OF THE INTERVENTION

The Productive Development Project (PDP) is one of three large-scale projects financed under the 2006 Compact between MCC and the Government of El Salvador.[1] The main objective of 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 project will use $87 million in allocated funds to benefit an estimated 55,000 beneficiaries over five years.[2,3]

PDP comprises three activities: (1) Productive and Business Services (PBS), (2) Investment Support, and (3) Financial Services. PBS activity is 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. Investment Support activity is designed to offer investment capital for competitively selected business proposals. Finally, Financial Services activity will provide technical assistance and financial resources to the banking sector and loan and output guarantees to small producers, as appropriate.

FOMILENIO will contract service providers to implement PBS activity. One overarching service provider, Chemonics, will coordinate and manage the activity's various components. In partnership with FOMILENIO, Banco Multisectorial de Inversiones (BMI) is likely to coordinate the Investment Support and Financial Services activities, although this may change in the near future.

MCC has contracted with Mathematica to design and conduct evaluations of the first two PDP activities—PBS and Investment Support activities. In principle, these two activities will be evaluated under different designs. However, the nature of these services allows for beneficiaries to receive a mix of the services offered under the PDP activities. To the extent possible, the 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 evaluation of the third activity, Financial Services, will be combined with the evaluation of the first activity, PBS, under the assumption that the financial services are just one more set of services eligible producers will be offered as part of the PDP intervention. This is

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

[2] The PDP will directly benefit an estimated 11,000 producers with technical and material assistance. Using an estimate of 5 persons per producer household, the PDP will benefit an estimated 55,000 people overall.

[3] Executive Summary, Millennium Challenge Compact with El Salvador, 2006.

particularly true if FOMILENIO decides to transfer responsibility for Financial Services activity from BMI to a consortium of financial entities led by Chemonics.

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

B. KEY RESEARCH QUESTIONS

The impact evaluation addresses the following primary research question: What impact did the offer of productive development services by FOMILENIO/MCC have on beneficiaries' incomes and employment?

Although assessing the impact on beneficiaries' income and employment represents the main goal of the evaluation, we will also address impacts on intermediate outcomes, such as business practices and plans, technology adoption, product diversification, and value chain integration.

Additionally, the evaluation is designed to examine the differential impact of offering PBS for two years instead of one year on beneficiaries' employment and income, as well as on the intermediate outcomes listed above.

C. PDP IMPACT EVALUATION DESIGNS

Given the diversity of productive sectors that PDP will target, Mathematica, MCC, and FOMILENIO have agreed that the impact evaluation should be limited to three sectors: (1) Artisan, (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. Thus, our preferred evaluation design is common for all three value chains but will be flexible enough to accommodate how the intervention will be implemented for each value chain. First, we present the core evaluation design and then we explain how this will be adapted for each value chain.

1. Design for PBS Activity

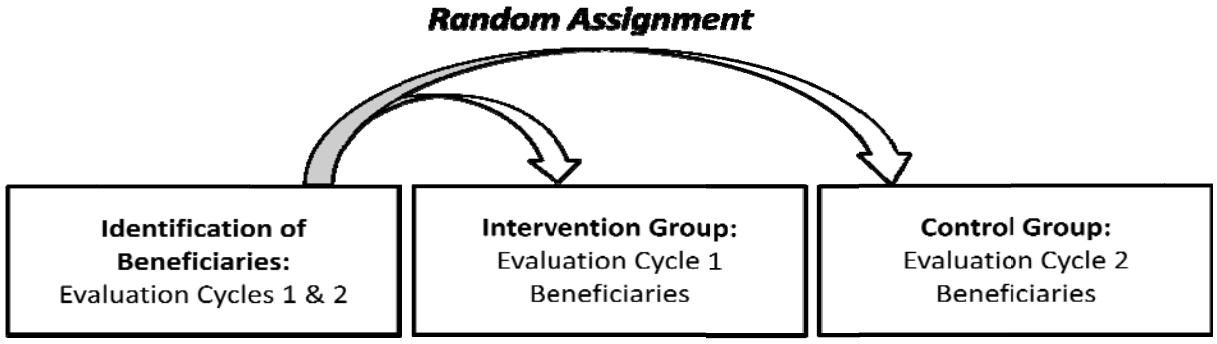
Our recommended design for evaluating PBS activity is a *randomized rollout design.* This design was accepted by all the stakeholders (MCC, FOMILENIO, and Chemonics, as documented in the attached agreement). It offers the key advantage of randomized studies: when implemented well, random assignment leads 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 proposed 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 will be 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 beneficiaries to be served in a subsequent cycle (the control group). Figure 1 illustrates the random assignment of beneficiaries into these two evaluation cycles. We will 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. However, we might need to augment the sample to address statistical considerations that arise due to the grouping or clustering of producers.

FIGURE 1

ILLUSTRATION OF CORE RANDOMIZED ROLLOUT DESIGN



The advantage of the randomized rollout design is that all eligible producers will be offered services and none will be turned down. The disadvantage is that impacts must be estimated before Evaluation Cycle 2 beneficiaries are offered the intervention services. The intervention cycle lengths for the three value chains to be evaluated are approximately 12 months for artisans and diary value chains and 9 months for horticulture. Thus, the design will not allow for estimating impacts that occur after one year for any of the three value chains, although we would be able to compare the effect of two years of the intervention with the effect of one year of the intervention.

Specifically, at the end of Evaluation Cycle 1, we will estimate the impact of approximately one year of program activities. 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, as noted, at the end of Evaluation Cycle 2, we will be able to compare the effect of two years of the intervention to the effect of one year of the intervention. At that point, the intervention group would have had access to two years of services and the control group would have had access to one year of services.

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, in most cases, Chemonics will offer services to groups of beneficiaries rather than individual beneficiaries. For example, the artisan intervention is being offered to groups of 10 to 15 beneficiaries. One of the principles on which the artisan intervention is based is attaining some degree of association within these groups in order to become more competitive in the artisan market. These beneficiary groups are defined in advance by the implementer, and all beneficiaries 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, a group of cantones or even a whole municipality—as opposed to offering the services across all cantones or municipalities. Because of these restrictions, random assignment will likely occur at the level of geographic clusters (for example, municipalities or groups of cantones.)

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 show up for 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.[4]

A disadvantage of randomly assigning groups or geographic clusters instead of individual producers is that larger samples will be 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

[4] 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.

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 only be able to detect large impacts. The section on estimating program impacts describes this problem in more detail.

An additional consideration is that because all eligible beneficiaries 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 only 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 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 would consist of the following eight steps (see Figure 2).[5]

1. Chemonics **identifies or recruits potential beneficiaries**. In this first and critical step, Chemonics recruits enough beneficiaries to fill the service slots available for Evaluation Cycle 1 and enough of Evaluation Cycle 2 to populate our 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 was agreed upon. In addition, the number of potential beneficiaries required for each value chain was agreed upon by Mathematica, MCC, FOMILENIO, and Chemonics based on Chemonics' target number of beneficiaries 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.[6] The list will also specify 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 beneficiaries 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

[5] 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.

[6] For the artisan value chain, the agreed sample size is 800 potential beneficiaries; for the diary value chain, 900 potential beneficiaries;, and for the horticulture value chain, 700 potential beneficiaries.

municipality) or in the same association. In this instance, we will 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 will need additional information about potential stratifying variables, such as the size of potential beneficiary groups, among others. These potential stratifying variables would most likely be specific to each value chain.

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, which will be served in Evaluation Cycle 2. The randomization procedure will take into account the exceptions and constraints discussed above and will be conducted separately within each of the strata identified. The evaluator will transmit the list of assigned potential beneficiaries to Chemonics on the agreed dates for each value chain in order for Chemonics to communicate to the beneficiaries whether they would be served right away or would have to wait for the next evaluation cycle.

4. **Collect baseline data** from all eligible beneficiaries before the start of Evaluation Cycle 1. The Dirección General de Estadística y Censos (DIGESTYC) will collect baseline data from all potential beneficiaries before each intervention starts. Although data collection could extend up to one month 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 will communicate 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).

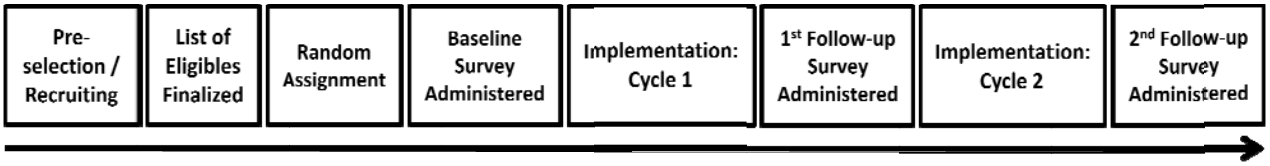
7. **Monitor the implementation of the intervention in Evaluation Cycle 2**. MCC and FOMILENIO expressed interest in assessing the impact of the intervention at the end of Evaluation Cycle 2. In this case, the evaluation will provide a comparison between receiving two years of the intervention services (group assigned to Evaluation Cycle

1) and receiving one year of the intervention services (group assigned to Evaluation Cycle 2). Mathematica and Chemonics will monitor the implementation during Evaluation Cycle 2 in order to identify potential problems and address them in a timely manner.

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 of the data collection section.

FIGURE 2

SEQUENCE OF ACTIVITIES OF THE CORE RANDOMIZED ROLLOUT DESIGN



2. Design for the Investment Support and Financial Services Activities

Our initial evaluation plan recommended using separate designs for the first two PDP activities—PBS and Investment Support activities. The proposed design for the Investment Support activity was a *regression discontinuity design*. Regression discontinuity exploits the availability of a continuous program selection score with an arbitrary cutoff point by examining the relationship between the score and program outcomes near the cutoff point. Because producers just below the cutoff are almost identical to those just above it—except that those above the cutoff qualify for the program—any differences between the two groups can be attributed to the intervention itself. Although regression discontinuity designs are capable of producing rigorous impact estimates, they often require substantially larger samples than designs involving the randomization of individuals and groups.

New information about the implementation of the Investment Support activity provided to Mathematica during our visit to El Salvador in May 2009 does not support the feasibility of using a regression discontinuity design for evaluating this activity. First, FOMILENIO and BMI informed us that the selection of beneficiaries would be done on a case-by-case basis and would not use a score with an arbitrary, specified cut point as originally planned. Second, the sample

sizes are likely to be considerably smaller than anticipated.[7] These two developments make a regression discontinuity design infeasible. Furthermore, the program might be redesigned shortly. Given these circumstances, Mathematica will reassess the feasibility of a rigorous evaluation of this activity after it has been redesigned.

The third PDP activity, Financial Services, will provide technical assistance and financial resources to financial organizations and loan and output guarantees to small producers. Although this study does not include a specific evaluation design for this activity, it is likely that the beneficiaries of the PBS activity will seek services under the Financial Services activity. Therefore, the evaluation of the PBS activity will take into account the possible overlap between the PBS and Financial Services activities. Our understanding is that the offer of Financial Services will not be tied in any way to the offer of PBS. Thus, Financial Services is simply another set of services available to both to the intervention and control groups.

The proposed evaluation design for PBS is, in principle, to evaluate only the offer of PBS. However, an unintended consequence of being offered PBS services could be to induce beneficiaries to request Financial Services. This would affect the interpretation of the impact estimates because the effect of PBS would be confounded with the effect of Financial Services. Because we cannot restrict the offer of services to the control group (that is, Evaluation Cycle 2 [control group] beneficiaries might request Financial Services before they are offered PBS), we cannot separate the impacts of PBS from the impacts of Financial Services under the proposed design. However, we can interpret the results of the evaluation in the context of a possible overlap of Financial Services and PBS. Mathematica will monitor the extent of overlap of Financial Services and PBS in the treatment and control groups and will use this information to interpret impact estimates appropriately.

Another project funded under PDP is the *Fund of Productive Initiatives*, recently launched by FOMILENIO and Chemonics. This project is not part of the evaluation, but we will monitor its activities as part of the context and alternatives available to the producers in the Northern Zone.

D. OUTCOME INDICATORS AND DATA SOURCES

The PDP impact evaluation will assess both main and secondary outcomes resulting from the offer of intervention activities. The main source for these outcomes is EDP, which will

[7] In July 2009, BMI informed us that only 13 proposals had been submitted for the Investment Support Activity of which only 2 were eligible.

consist of customized survey instruments for each of three value chains featured in the impact evaluation.

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 artisan, dairy, and horticulture production. We will also track employment information, measured by the number of days worked in the past cycle. Table 1 summarizes the evaluation's two main outcomes and their corresponding indicators.[8] 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 adoption of new practices and technologies, as well as enhanced product diversification and value chain integration. Table 1 summarizes the evaluation's four 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.

TABLE 1

KEY PDP MAIN AND SECONDARY OUTCOMES

Outcome	Indicator	Time of Collection
Main		
Employment	Number of days worked in past cycle	Baseline, first, and second follow-ups
Income	Household income in past cycle, including income from productive development	Baseline, first, and second follow-ups
Secondary		
Business Practices and Plan	Group has composed a business plan and used basic business practices in past cycle	Baseline, first, and second follow-ups
Technology Adoption	Group has used new technologies and the	Baseline, first, and second

[8] 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 artisan and dairy value chains, one evaluation cycle is 12 months long. For the horticulture value chain, one evaluation cycle is 9 months.

Outcome	Indicator	Time of Collection
	internet to acquire materials and produce/sell products in past cycle	follow-ups
Diversification	Group produced a large number of products with a variety of raw materials in past cycle	Baseline, first, and second follow-ups
Value Chain Integration	Group had multiple sources of information regarding prices and preferences, as well as formal contracts with buyers in 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[9] on employment and income, we must collect primary data on baseline characteristics, outcomes, and utilization of PBS for beneficiaries in the study sample. We will collect these primary data through EDP, which are surveys developed specifically for this impact evaluation. EDP will be administered to all eligible beneficiaries in the three value chains featured in the evaluation; this includes all beneficiaries in the intervention group as well as the control group.

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

a. Survey Instruments. Both baseline and follow-up versions of EDP will capture data on the key outcomes mentioned above; baseline surveys will measure beneficiaries' initial practices, employment, and income, whereas follow-up surveys will monitor how beneficiaries' practices, employment, and income change throughout the study period. In addition, EDP will also collect background and participation data. Background data include demographic information about individuals and their communities. Participation data will detail beneficiaries' 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 beneficiaries' characteristics, resources, and productive activities over the course of the study.

[9] If Financial Services are tied to PBS, we will also evaluate the combined effect of the offer of PBS and Financial Services.

Within each productive chain, DIGESTYC will administer two distinct survey instruments: an individual instrument and a group leader instrument.[10] All beneficiaries 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.

TABLE 2

TOPICS COVERED BY EDP

Individual Instrument	Group Leader Instrument
General Demographic Information	General Group Information
Household Roster and Characteristics	Group Production
Artisan Activities	Group Sales
Artisan Costs and Income	Common Problems Among Members
Household Expenses and Income	Collective Productive Activities
Credit	Group Costs and Income

b. Survey Sample Frames and Sampling Plan. The target populations for each of the three value chains are eligible beneficiaries as determined by Chemonics. In the artisan chain, these eligible beneficiaries are organized in productive groups. In the dairy and horticulture chains, some eligible beneficiaries are members of productive groups, but others are not. In all chains, Chemonics will identify beneficiaries through a formal recruitment process for each value chain. First, Chemonics will hold a series of meetings with artisans, dairy producers, and farmers in various municipalities in the Northern Zone and administer a simple questionnaire to all interested producers. Second, Chemonics staff will compile a master list of all interested *and* eligible beneficiaries for each value chain. These master lists comprise the complete sample frames for 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, EDP will yield a longitudinal data set of PDP beneficiaries and

[10] We anticipate developing an individual and leader survey for each of the study's three value chains, for a total of six unique survey instruments.

productive groups, although there might be cases in which beneficiaries drop out from the intervention and cannot be located or refuse to respond to the follow-up surveys.

The municipalities represented in master lists will vary based on demand for each value chain's services. As of this writing, an estimated 24 municipalities will be represented in the master list of eligible beneficiaries for the artisan value chain. We estimate that multiple municipalities will be represented in master lists for the horticulture and dairy value chains.

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, 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.[11] 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

[11] As of September 2009, DIGESTYC and FOMILENIO have signed only one contract regarding the administration of the baseline EDP for the artisan chain. However, future contracts will designate DIGESTYC's responsibilities in administering baseline EDPs for the other two value chains, as well as follow-up surveys for all three value chains.

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 beneficiaries' cooperatives, workshops and groups to survey all beneficiaries in the sample frames.[12] 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 beneficiaries will be interviewed during these meetings. Following the meetings, DIGESTYC surveyors will contact and/or travel to the homes of beneficiaries 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 beneficiaries to speak with other members of the beneficiaries' 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.

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 beneficiaries 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 artisan value chain cycle begins prior to the dairy and horticulture chains, DIGESTYC will administer the artisan baseline survey from October to November 2009. The horticulture baseline survey will be administered from April to May 2010; the dairy survey will be administered from March to April 2010. Follow-up surveys for the artisan and dairy chains will be administered on a 12-month cycle; follow-up surveys for the horticulture chains will be administered on a 9-month cycle. Table 3 outlines additional key dates related to PDP data collection, including dates of follow-up surveys.

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

TABLE 3

DATA COLLECTION PERIODS, BY VALUE CHAIN

	Value Chain		
	Artisan	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	March 2010	March 2010
5. Baseline data collection	October - November 2009	April - May 2010	March - April 2010
6. Baseline data set sent by DIGESTYC to Mathematica	January 2010	July 2010	June 2010
7. Data review	February 2010	August 2010	July 2010
8. Baseline data collection documentation sent by DIGESTYC to Mathematica	February 2010	August 2010	July 2010
First Follow-up Survey			
9. Develop draft follow-up survey instrument	February - April 2010	August - September 2010	October - November 2010
10. Revise instrument based on comments from Chemonics and FOMILENIO	May 2010	October 2010	December 2010
11. Conduct interviewer training	July 2010	December 2010	February 2011
12. Follow-up (Round 1) data collection	August - September 2010	January - February 2011	March - April 2011
Second Follow-up Survey			
13. Develop draft follow-up survey instrument	April 2011	June 2012	November 2011
14. Revise instrument based on comments from Chemonics and FOMILENIO	May 2011	July 2012	December 2011
15. Conduct interviewer training	July 2011	September 2012	February 2012
16. Follow-up (Round 2) data collection starts	August - September 2011	October - November 2012	March - April 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 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 alleviate 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 beneficiaries in each value chain; the case of individual randomization is a special case of this general model. 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 error term (a group or cluster "random 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, and so on. 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.

The core model can be modified to explore alternative specifications. For the case of individual random assignment, the treatment indicator will have the subscript i. The subscript g will be eliminated from Equation (1); the group or cluster random effect will also be deleted

from the equation. The specification presented in Equation (1) can also be simplified to exclude a group-specific random effect, which will make it a simpler regression in which we adjust the estimated standard errors for group clustering (that is, we would use the Huber-White estimator for estimating the variance of λ). The advantage of specification (1) over the simpler regression with a clustering correction is that the variance of the impact estimate is likely to be smaller. However, the simpler regression could accommodate a combination of individual random assignment and group random assignment and it will be easier to estimate. We will assess the sensitivity of our impact estimates to model specification when the data are available.

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, which further reduces the intervention's homogeneity across chains. Therefore, our recommendation is to focus on obtaining precise impacts for each value chain. 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.

2. Statistical Power

We have made preliminary calculations about the required total sample size needed to detect the target impact estimates for each value chain. We conducted our calculations using the log of individual agricultural income from the 2007 Encuesta de Hogares de Propósitos Múltipes (EHPM) because its variance was substantially smaller than the variance of the untransformed income, which resulted in smaller minimum detectable effects (MDEs) and, therefore, smaller sample sizes. For these calculations, we made several assumptions: (1) an intra-class correlation for individual log income from agricultural activities estimated at the cantón level from the 2007 EHPM equal to 0.026; (2) a total attrition rate of 15 percent for both groups and individuals; and (3) the percentage of variance explained by the regression model (R^2) equal to 0.5 for both groups and individuals.[13]

[13] We are also assuming 80 percent power, 5 percent significance, and a two-tail test.

Table 4 presents the effect sizes and net impacts that the study will be likely to detect given a target sample size. As noted above, the target sample sizes agreed upon by all the stakeholders were 800 potential beneficiaries for artisans, 700 for diary, and 900 for horticulture. Given these sample sizes, we present the estimated size of the effects we are likely to detect: (1) the MDE, which is measured in standard deviations; and (2) the equivalent minimum detectable impact (MDI), which is measured as the percentage increase in household income at the end of the intervention relative to the control group. In addition, we present two possible scenarios of random assignment: (1) random assignment of small groups, which could be thought of as random assignment of groups of cantones; and (2) random assignment of large groups, which could be thought as random assignment of municipalities. The tradeoff is that when randomizing larger groups (municipalities), the MDEs are larger, all other things being equal.

To facilitate the presentation, we refer to the case of small groups (groups of around 14 to 17 members) as assignment of groups of cantones, and the case of large groups (groups of around 27 to 35 members) as assignment of municipalities. As the benchmark case, we use the former.

For the artisan chain, if 800 producers are distributed in groups of cantones with approximately 15 producers, then the study will randomly assign 44 groups and we will be able to detect an income change of 0.18 of a standard deviation, which translates to a 37 percent change in income. For the case of random assignment at the municipality level, assuming that the 800 producers are distributed in municipalities of approximately 30 members, then the study will randomly assign 22 municipalities and will be able to detect an income change of 0.20 of a standard deviation, which translates to a 42 percent change in income.

For the dairy value chain, if 700 producers are distributed in groups of cantones with approximately 14 producers, then the study will randomly assign 42 groups and we will be able to detect an income change of 0.19 standard deviations, or equivalently, a change in income of 39 percent. For the case of random assignment at the municipality level, assuming that the 700 individuals are distributed in 22 municipalities with approximately 27 producers, we estimate that we can detect an income change of 0.21 standard deviations, or equivalently, a 43 percent change in income.

For the horticulture value chain, if 900 producers are distributed in groups of cantones with approximately 17 producers, then the study will randomly assign 44 groups and we will be able to detect an income change of 0.17 standard deviations, or equivalently, a 35 percent change in income. For the case of random assignment at the municipality level, we assume that the 900 producers are grouped in 22 municipalities with approximately 35 producers, we estimate that

we can detect an income change of 0.20 standard deviations, or equivalently, a 40 percent change in income.[14]

TABLE 4

ESTIMATED DETECTABLE IMPACTS ON THE LOG OF INDIVIDUAL AGRICULTURAL INCOME FOR THE TARGET SAMPLE SIZES FOR EACH VALUE CHAIN, UNDER DIFFERENT CLUSTERING SCENARIOS

Random Assignment Scenario	MDE	MDI (Percentage Change in Income)
Artisans		
Random assignment of 44 groups of cantones with 15 individuals	0.18	37
Random assignment of 22 municipalities with 30 individuals	0.20	42
Dairy		
Random assignment of 42 groups of cantones with 14 individuals	0.19	39
Random assignment of 22 municipalities with 27 individuals	0.21	43
Horticulture		
Random assignment of 44 groups of cantones with 17 individuals	0.17	35
Random assignment of 22 municipalities with 35 individuals	0.20	40

Source: Mathematica calculations based on data from the 2007 Encuesta de Hogares de Propósitos Múltiples (EHPM) and from the 2005 Censo Agropecuario de El Salvador.

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

[14] Note, these calculations assume that all groups have the same number of members (balanced groups), if in reality the groups greatly differ on the number of members then the detectable impacts will be even lower.

F. REPORTING PLANS

We will provide short memoranda summarizing the findings at three points in the lifecycle of each value chain. Baseline memoranda will summarize findings from the baseline EDPs and will analyze the characteristics of the intervention group versus the control group for each of the three value chains. First follow-up memoranda 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 these reports is to quantify the impact of one cycle of productive development services on beneficiaries' incomes and employment and other outcomes. Second follow-up memoranda 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. The main focus of these final memoranda is to quantify the impact of two cycles of productive development services versus one cycle of services.[15] Table 5 presents tentative dates of all key deliverables associated with the PDP impact evaluation.

TABLE 5

KEY DELIVERABLE DATES, BY VALUE CHAIN AND DATE DUE

	Productive Chain		
Deliverable	Artisan	Horticulture	Dairy
Baseline Analysis	February 2010	September 2010	August 2010
Impact Analysis (First Follow-Up)	January 2011	June 2011	August 2011
Impact Analysis (Second Follow-Up)	January 2012	March 2012	August 2012

cc: Liza Valenzuela (MCC), Damiana Astudillo (MCC), Ricardo Orellana (FOMILENIO), Carmen Valle (Chemonics), Francisco Munguía (DIGESTYC), Miguel Montesino Hernández (AENOR Centroamérica), M. Induni, File

[15] As noted, we would write nine memoranda over the course of the evaluation. Another option is to combine findings from the three value chains at each lifecycle stage to produce one baseline report, one first follow-up report, and one second follow-up report over the course of the evaluation. We will discuss these, and other, reporting options with MCC in future weeks.