Committee On Sustainability Assessment

COSA Methods

Concise introduction to the methodological principles

COSA Methods

These methodological principles compiled here serve as a concise review of the most important elements of the Impact Assessment and Performance Monitoring that are supported by the Committee on Sustainability Assessment (COSA). This brief overview explicitly outlines the steps and processes we undertake. For more information on the underlying principles or about how we work with Impacts, Impact Pathways, and Theory of Change see {The Principles and Characteristics of the COSA System}, a companion document that is best read first to better understand the context of these methods.

Contents

Systematic Assessment of Sustainability

COSA is not simply a research organization, we are mandated to merge scientific rigor with business-like pragmatism that is necessary for strategic and common sense decision-making about agricultural sustainability. COSA has a well-established framework for performance monitoring and impact evaluation.

Our approach to more effective measurement stems from our commitment that being effective in sustainability is not just about being practical, it also manifests as core COSA principles that maximize value for all stakeholders. These include:

1. Being **rigorous**, regardless of methodologies used
2. Being **respectful** towards the people and place we evaluate or research. These are not subjects in the process but rather participants
3. **Measuring only what matters** for decision-making

These values relate to ideas of "Impact Investing", "Lean Research", "Client-Centric", and "Applied or Bottom-up" research but go beyond concepts to the use of tested tools and modules that achieve results. The COSA System is built around three major steps for measuring and understanding sustainability. The core process characteristics are gathered under these themes:

1. **Define the Pathway** – Working closely with clients and partners to develop a realistic understanding and practical pathway to sustainability
2. **Gather the Facts** - Ensuring you have the right info at the right time
3. **Answer & Advise** – Employing broad experience and state-of-the-art analysis to get the most out of the data

Figure 1: The COSA System and its Main Components

Define the Pathway
Developing a realistic understanding and practical pathway to sustainability

- Aligning with international norms
- SMART standardized indicators
- Multi-dimensional framework: environmental, social and economic

Gather the Facts
Ensuring you have the right info at the right time

- Vetted Surveys
- Digital data capture and geospatial mapping
- Building local capacity and global standardization
- Performance monitoring tools

Answer and Advise
Employing broad experience and state-of-the-art analysis to ensure you get the most out of the data

- Credible scientific analysis
- Knowledge base and benchmarking
- Customized KPI management dashboards
- Better informed policies

The components of the COSA system are illustrated and expanded below

1. Define the Pathway

The Committee on Sustainability Assessment has designed steps for getting and using information to improve sustainability management practices. For clients, projects, and firms we guide a process to help clearly define the purpose of this work and to map the existing internal sustainability framework e.g. “What questions do you want to answer?” and “What will you do with the results?” This then guides the choice of focal areas and target issues to be addressed and how to best do that. COSA provides guidance to the approaches that are most appropriate and effective for both Partners and clients to achieve their objectives. This includes support in selecting the optimal indicators and methods that will illuminate an appropriate sustainability pathway.

Alignment with International Norms

To help ensure transparency and global acceptance, COSA has systematically aligned its Indicators and work with dozens of important multilateral and multi-stakeholder instruments, ranging from the Bellagio Sustainability Assessment and Measurement Principles and the Rio Declaration to the International Labor Organization’s Core 8 Labor Standards and the OECD Economic Guidelines.

Selecting the Right Indicators

COSA and dozens of partner institutions are adopting a harmonized framework of globally-accepted indicators. These are validated by diverse experts and tested in many thousands of surveys.

The COSA indicators underwent a painstaking evaluation of whether or not they met SMART characteristics (Specific, Measurable, Attainable, Relevant and Time-bound). COSA indicators are also:

- Aligned with key international accords
- Generally comparable across different conditions, crops, and situations
- Oriented to measurably change over the short to mid-term
- Sufficiently specific in definition to ensure clarity and comparability to ensure that the same thing is measured each time in the same way
- Measurable with reasonable cost and effort

These best-practices ensure sound data and a common language for sharing information and comparison.

There are also a number of global themes or categories that we can readily calculate and present in our work as indices and cross tabulations where these are valid. A prominent example is our testing and application of the Progress out of Poverty Index (PPI) – now used in more than 40 countries – in collaboration with the Grameen Foundation.

What to Observe: SMART Indicators

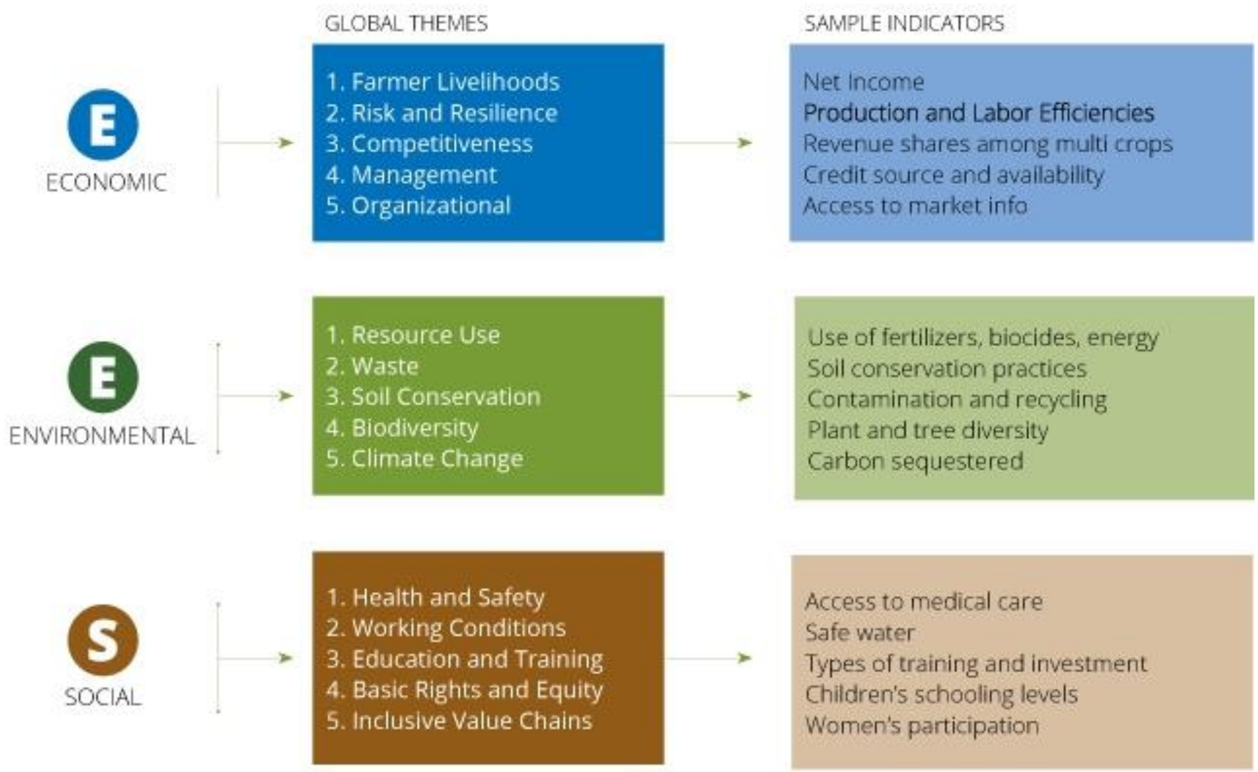
Good surveys use thoughtful SMART indicators (see box on this page) as a starting point. This helps to establish a common framework for the definition and collection of comparable sustainability data so that learning is facilitated (see Fig 2). COSA Indicators are adapted to the crop type (e.g. coffee, cocoa, food crops) and are sufficiently standardized to allow comparison across countries, projects, and even crop types (for many of the measures such as comparative net income, costs, etc.).

Importance of multi-dimensionality

Simple assessments or views run a high risk of missing key factors that compromise projects, investments, and reputation. Although it is easy and tempting to just measure one or a few factors such as farm yields, income, or biodiversity as the proxy for sustainability, the reality is that sustainability, by definition, necessitates balancing social, environmental, and economic facets. Any measurement that does not take this holistic view into account is simply not assessing sustainability. For example, if higher yields are achieved by clear-cutting forested areas, which then results in soil erosion, silted waterways, and the loss of timber and firewood for the surrounding communities, it may not be a sustainable outcome. This can present quite a challenge for projects or investments whose focus is limited to only one or two desired outcomes so managing the unexpected outcomes is important.

COSA engages a variety of methods, besides the balanced use of SMART indicators, to observe the different dimensions of sustainability ranging from simple cross-tabulations of data points to stochastic frontier analysis and to relational analysis with, for example, the integration of the Progress out of Poverty Index (PPI) or the Multidimensional Poverty Index (MPI).

Figure 2: Sampling of Global Themes that Inform Major COSA Indicators



Defining impact Asssessment

Impact Assessment is simply defined as the intended or unintended longer-term effects (both positive and negative) that can be attributed to a specific intervention or investment. This can include aspects such as competitiveness, ecosystem health, or consistently different income levels. COSA further distinguishes its impact assessment by inclusively looking beyond single dimensions to include the environmental, social, and economic manifestations of change so as to better understand the relation of any impact to balanced sustainability.

Activities, such as financing or training can lead to an impact but are considered as interventions or outputs. If an expected impact such as yield increase occurs after training on productivity, then the causal pathway is more clear and the results are more likely attributable to the training or intervention, assuming that similar untrained farmers did not have similar yield results. Knowing the likely counterfactual or what occurred for similar farmers in the absence of the intervention is a key aspect of credible impact assessment.

2. Gather the Facts

Building local capacity

We start by building local capacity with capable partner institutions in each target country to achieve a better scientific process that features local relevance and a richer contextual understanding.

Adaptation: balancing local relevence with global standardization

An initial stakeholder workshop and a pre-research review are designed to collect information from different sources and a broad spectrum of participants and local experts. This information, gathered in a systematic way (using standard Country Conversion and Context Worksheets) enables adaptation to local conditions.

The Context Worksheet includes important questions about the project goals, proposed interventions, theory of change, and important specifics of the sampling areas. It facilitates good project structuring and also provides insights in the eventual analytical approach to potential challenges or opportunities. For example, we may learn of state-sponsored subsidies or training in the same sampling areas that could spill over into our control groups.

Representative Samples of Producers and Communities

Sample design starts with selecting representative farms (such as those that undergo an intervention, adopt sustainability measures, or take part in a supply chain project). For impact assessment, the next step is to select control farms that are similar in many obvious and less obvious ways. It is vital to capture the same key criteria that drove the selection of the target group and/or other factors that are likely to influence their outcomes or performance.[1]

Within this basic with/without design, techniques such as stratification or clustering - particularly at village and organizational levels - can be used. The samples are selected to allow both descriptive analysis and econometric analysis that detect, with high levels of confidence, the differences in the performance between target and control farms.[2]

Whether sampling at the household or cluster level, it is important to randomize among the treatment, so that with a large enough sample, the only relevant difference between target and control groups is the treatment itself. Even with randomization and thoughtful selection, there will likely be a degree of self-selection bias when participation in the treatment is voluntary. Aside from setting up a Randomized Control Trial (RCT), attempting to correct for self-selection can also occur ex-post during the analysis.

Regarding the sample size, the general guideline, besides financial constraints, is sufficient representatives. The formal way to determine sample size is through a power calculation, which would use information gathered from the Context Worksheets to make sample size decisions. For example, if a client knows specifically that they hope to see a 10 percent increase in yield for target producers, we can potentially run a power calculation to determine the necessary sample size to have a good chance of capturing that change. Software like Optimal Design offers easy layouts and input boxes to calculate power and sample size.

[1] For example, two 5 hectare farms might differ considerably in key aspects such as: asset levels, agro-ecological conditions, agri-business organization, infrastructure, entrepreneurial capacity, etc.
[2] Optimal levels of statistical confidence (e.g. 99%) are not always viable but COSA typically reports the levels.

Respecting the Counterfactual: How to select a control group

In order to understand the counterfactual (or what would have happened in the absence of an intervention or investment) COSA will often simultaneously measure control groups as well the target or treatment group. These control groups can only be considered valid to the extent that they are functionally similar to the target farmers and differing primarily by not having the same investment or intervention such as certification, credit, or training. Control farmers can be selected on the basis of their farm size, experience, Agro-Ecological Zone, membership in cooperatives or associations, distance to markets, level of assets, ethnicity, and more. There are considerable challenges to finding appropriate controls that are not always obvious. For example, farmers may benefit from access to diverse types of infrastructure, they may have other training or certifications, they may receive unseen revenue, and they may have diverse motives or entrepreneurial capacity. Active due diligence is necessary in order to identify good control groups and manage potential bias.

Furthermore, anticipating the analytical strategy is important in choosing a control group. For example, in order to employ a Propensity Score Matching (PSM) analysis it is necessary to ensure a large sample of control farmers to serve as appropriate matches for the target producers (see below). In cases when finding an adequate number of controls is impossible, we would eliminate PSM from the options.

Surveyor Selection and Training

Field surveyors (enumerators) are selected for a balance of local knowledge, practical understanding of survey work and interest in sustainability principles. While all surveyors are trained for about a week before the project, it will be necessary to continually train, correct, and motivate surveyors, since much of the work takes place in difficult rural conditions and requires randomly checking surveys and consistently observing surveyors in the field. We also train surveyors in advanced technology, such as using smart phones or tablets with built-in data validation to ensure reliable data. {see Training Protocol}

Piloting the Survey

In addition to the initial context analysis and review with key experts-stakeholders, there is an opportunity to make the necessary adjustments to surveys after the initial pilot surveys. Modifications are made where needed to ensure consistency of survey meaning and to allow for more accurate results that can be presented in a globally consistent manner. Along with testing the survey, the pilot is also an extension of surveyor training and an opportunity to understand potential logistical challenges, such as transportation or internet connectivity.

Given all the opportunities to learn from the pilot, it is important to build in time (several days) between the pilot and the actual survey in order to make changes to the survey, further train surveyors, and prepare any necessary logistical changes.

Vetted Survey & Field Tools

Data is collected primarily via structured digital surveys conducted by trained local professional surveyors. On average, surveys take 1 to 1.5 hours and include both direct observations and structured questions.[3] We also conduct surveys of key producer organizations or enterprises since these can have a considerable effect on the impacts. For Performance Monitoring, a related

[3] Surveys of Producer Organizations are separate and distinct and can take longer to conduct {see PO Survey}

but much simpler tool is developed to closely serve the managerial needs of a project or investment and it can be used with minimal training.

Digital data capture

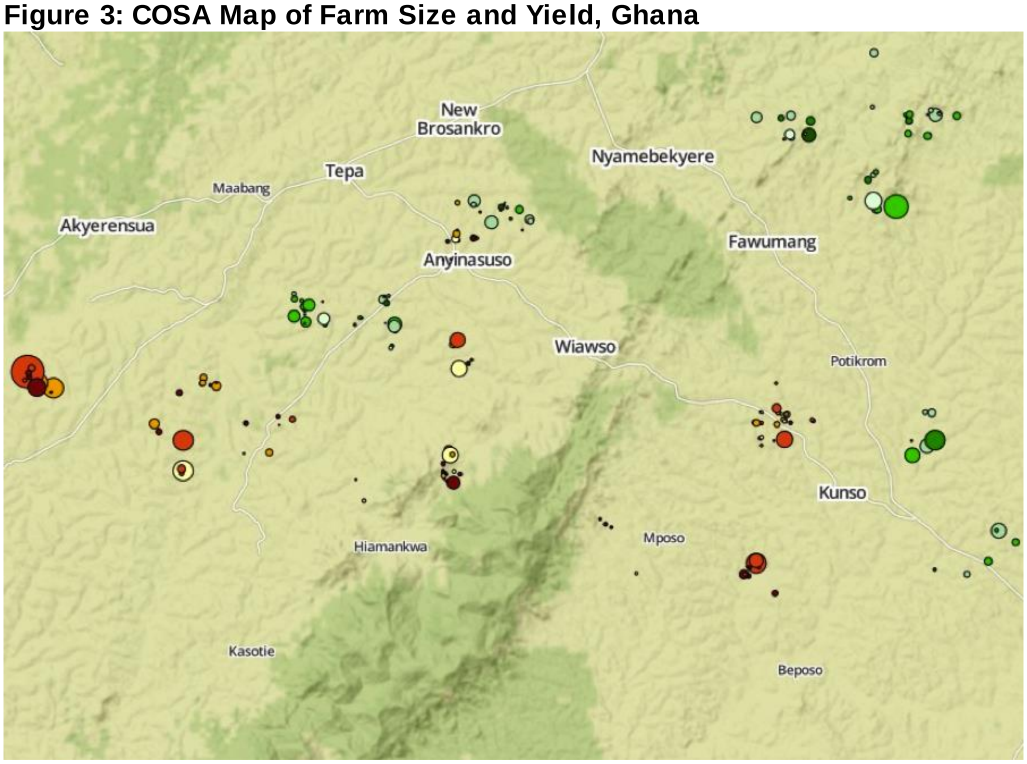
COSA uses diverse technologies to accelerate understanding in sustainability. For example, our COSATouch surveys have skip-logic, internal answer validation checks, and offline functionality. Built-in validations reduce input errors to improve accuracy and reduce data cleaning. Surveyors can use any sort of hand-held devices (Android or Apple tablets, laptops, smart phones). These features combine to make the process of interviewing farmers and group leaders speedier and more conversational than would be possible with paper and a clipboard.

Another advantage of using COSATouch is the ability to automatically capture place and length of survey time that, in combination with remote monitoring of responses, allows real-time quality control by managers.

Geospatial mapping

Geospatial mapping becomes particularly important as a tool to understand the geographic implications of the array of indicators. Mapping data geospatially provides a more systemic view to refine diagnosis of issues in the field and to incorporate and test solutions where regional variance is likely. Seeing the geophysical attributes of the regions we study allows us to develop different hypotheses about why one region is doing better than another and whether low-performing regions might benefit from improved targeting or a different type of intervention. Opportunities exist to also integrate farm-level data to the remote sensing of land-use changes to ascertain the correlations between certain socio-economic factors and changes in the environment.

Figure 3: COSA Map of Farm Size and Yield, Ghana



Producer Organizations

We also take into account key institutional factors such as the influence of Producer Organizations (PO). Our globally vetted Producer Organization Survey enables us to understand their membership, structure, capacity, governance, assets, and services since POs can be more influential than many other factors.

Qualitative Methods

We most often apply a mixed-method design with an appropriate mix of quantitative and qualitative tools selected to best capture and then focus data to inform decision making. This typically includes contextual assessments and perception questions. In conjunction with the primarily quantitative standard that we commonly apply, researchers may propose those qualitative methodologies that may be appropriate. A sampling of these can include: Contribution Analysis, Most Significant Change, Farmer Portraits, elements of the 5 Capitals approach, gender and crop calendars, livelihood zone mapping, Sensemaker, participatory film and video.

Performance Monitoring

The COSA Performance Monitoring System is a fast and affordable management tool for measuring sustainability performance. It provides instant feedback to managers of sustainability projects and investments on a range of indicators and mission-critical questions. This customized data is reported in a real-time dashboard to facilitate day-to-day decision making.

Performance Monitoring is easily integrated into normal business operations to capture KPI — surveys are administered during normal field operations by local staff or technical advisors, and require only minimal training. Surveys are brief (typically 10-15 minutes) and are conducted on a regular basis, allowing managers to continuously keep projects on track.

COSA's Performance Monitoring System can be a stand-alone tool or it can also be paired with COSA's more sophisticated baseline and Impact Assessment tools to improve the accuracy and credibility of results.

3. Answer & Advise

Scientific credibility starts with understanding the context and gathering the right data. Only then can the data be mined with a broad array of thoughtful diagnostic and analytic tools. COSA is interested beyond simply knowing something, it is also important to help identify attribution and the reasons for an outcome. This permits a practical solution or the better focus of an investment or policy.

Basic Diagnostics

COSA's impact assessments use a mixed-method approach that better captures and assesses the diverse conditions found in the field. While basic scientific principles must underlie all sustainability analyses, needs and perspectives vary. The main component of the approach is the use of two standardized surveys: one administered to farmers, and another conducted with cooperatives or the community level organization that interacts with farmers. This process is informed and bolstered by the integration of useful secondary data gathered from field visits and key stakeholders prior to the assessment. Preparation is the key to a good survey and, like preparing for a sporting event, it must be done diligently and takes much longer than the actual event itself.

COSA makes considerable use of primary demographic data to understand how factors such as *farmer age, gender, education,* and *relative wealth* could influence outcomes, as seen in Table 4.1. Among other key variables that are taken into account are the *types and quantity of training received*, the *Agro-Ecological Zone* in which they operate (soils, slope, precipitation, etc.), *information they have access to, recent shocks* (civil or climactic disturbance), and the *distance to markets*. All of these can influence the outcomes and are important in order to understand and account for differences that may exist between the households or farms so that the results of an intervention can be tracked more accurately.

Table 4.1 Key Characteristics

Household Demographics	Producer characteristics	Age of decision maker (producer) responsible for the focus crop, grades of school completed, gender, years of experience growing focus crop
	Household revenue	Combined revenue from focus crop sales, other crops, other earnings (off farm employment, services, business revenue, land or equipment rental), and gifts & remittances
	Household composition	Number of people, genders, ages, dependency ratio, literacy, and school grades completed
Farm Characteristics	Farm characteristics	Management by owner (renter or sharecropper) or by a paid manager, farm size, focus crop area, farm location (GIS coordinates), distance from farm to nearest commercial center
	Land tenure	Owned by farmer, rented, sharecrop, communal ownership, farmed without payment
Change events	Shocks	Occurrence of major events that led to a significant change in the household's income, assets, or consumption.

Along with providing valuable context for understanding and structuring a proper assessment as well as interpreting outcomes, the demographic variables can also be used to improve the analytic approach. For example, in the case of a PSM analysis, we use many of these demographic variables to match target and control producers to account for a combination of differences and allow optimal scientifically-matched comparisons.

A number of these variables and other relevant information are collected during the project set-up stage through a questionnaire that has been refined over time. This guides the way the data collection is designed as well as how it is later analyzed and reported. We also focus on the context with local stakeholders in order to have much more refined and nuanced understanding of the realities that farmers face, and to help identify the likely pathways or approaches to best work with them.

Data Cleaning

Thoughtful and well-adapted surveys, using real-time validations, and the training and monitoring surveyors all contribute to accurate data. But there will always be a need to clean the data before performing the analysis. While too extensive to detail in this basic document, the core guidelines are:

1. Ensure that all the questions asked in the survey appear in the dataset.
2. Check that missing values result from skipped responses and should not be a value of zero, and vice versa.
3. Perform simple outlier analyses, especially on key variables such as farm area, production, trees or plants, labor days, etc.
4. Document all cleaning choices and tag data appropriately

It is important to remain in close contact with field managers, surveyors, and others involved in the data collection process who have valuable contextual knowledge of why questions were answered a certain way. Even within the same country, regional or cultural differences may cause producers to understand questions differently.

Managing Bias

COSA's field experience helps to design sampling and analytical techniques that minimize important biases. For example: controlling for sampling bias through Propensity Score Matching (see Analysis section below); controlling for spill-over effects by selecting control groups from separate but similar communities; controlling some of the self-selection bias through the context assessments and random selection of target and control groups; and controlling for institutional differences between the two groups.

To date, COSA has focused on observational studies, but it plans to include other forms of investigation including randomized control trials in the near future. Randomized control trials (RCTs), when well-constructed, allow somewhat more confidence in assuming the causality of observed outcomes to an intervention. While they can at times be narrowly focused, they can nonetheless serve as one of the complementary methods COSA uses for developing a rigorous and balanced understanding of the challenges and dynamics of agriculture systems.

Credible Scientific Analysis

In its analytic work, COSA has successfully utilized several approaches to better qualify and understand the available data. While most impact assessments use only a two-period model (baseline and follow-up), the results are more rigorous and insightful in multi-period models.

Sustainability is intrinsically complex and multi-dimensional. Some of the techniques we use to understand this complexity and communicate it in a straightforward way are:

- Cross-tabulations to look at the interaction between two or more indicators. For example, we may look at yield levels tabulated together with the use of fertilizer and with participation in training to assess what the correlation may be between productivity and the use of physical inputs or improved capabilities. We might further observe how this changes by gender or level of education or region.
- Efficiency analysis, production cost analysis, or stochastic frontier analysis.
- Comparing sample means of indicators to regional or country means – this practice offers perspective to our findings. Similarly, understanding the distribution curves of data suggests whether it is best represented as a mean or otherwise.

- Using the Progress out of Poverty Index (PPI) or the Multidimensional Poverty Index (MPI) as proxies for economic and social well-being to better understand the context for the findings.

The best analysis is not limited to one or another approach. COSA openly engages a broad toolkit of options and can integrate any credible approach to strengthen its analysis. A sampling of the methods includes:

1. *Difference in Differences (DID)* compares, using a simple linear model, the difference between the groups at baseline with the difference realized between the control and target after the intervention. Using this control group as a comparison at baseline helps control for differences between groups and helps mitigate the impact of how variability in conditions (independent of those caused by the intervention) may affect many of the observed changes. This is especially the case in agriculture, where yields (for example) can be significantly affected by local phenomena that can vary substantially from year to year.
2. *Propensity Score Matching (PSM)* is a statistical matching technique used to more accurately compare groups by estimating the effect of a policy or intervention (treatment) by accounting for factors that may predict receiving it and could affect indicator performance. PSM helps address self-selection bias wherein producers choosing certification may be intrinsically different from producers who do not (e.g. they may be more entrepreneurial, higher-yielding, or have more access to credit).
    - PSM is data intensive, as it requires many control producers to serve as potential matches for target producers. This technique will drop some producers from the sample if it cannot find a suitable match, depending on a pre-defined threshold. For more on how to define this threshold see: http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=35320229
    - PSM also requires a range of demographic and contextual variables, gathered from the producer surveys and potentially from the PO survey. It is important to collect this data and allow the analyst to select the appropriate variables by which to perform the matching.
3. *Instrumental Variables (IV)* analysis is used to estimate a relationship when there is simultaneity (when the casual direction of a treatment is not immediately obvious rendering the results of simple regression analysis biased and inconsistent). This method “replaces” the endogenous variable (the treatment variable) for a highly correlated but exogenous proxy. In this case, the IV approach uses the exogenous variables, (instruments) that predict treatment but do not predict the outcome variables (our indicators).
    - When appropriate, using IV is an effective and widely-accepted way to establish attribution of the outcomes to the intervention. Any variable that impacts outcome variables by way of the treatment can be used; yet identifying a good instrument can be more of an art than a science – finding a good exogenous proxy requires good background or contextual knowledge of the project.
4. *Stochastic Frontier Analysis (SFA)* uses the measured yields and inputs to estimate the highest level of yield that can be achieved for that sample of producers given the inputs utilized. It estimates the level of inefficiency for producers who did not reach that level and can estimate the components that might have contributed to this level of inefficiency. Because not all input data is available or relevant in each area of study, each SFA has a

slightly different specification for the stochastic production function, though all the pertinent inputs are included. We follow the conventional specification for SFA and a simultaneous equation to explain the inefficiency term using components relevant to input use, such as producer demographics (sex, age, education), input technology (use of equipment) and locational fixed effects.

5. *Regression Discontinuity Design* is a technique applicable when a cut-off point on a continuous variable (such as a poverty index or yield cut-off) is used to determine who receives a given intervention or project. The impact of the intervention can then be estimated by comparing outcomes for producers whose scores just qualify for the project with outcomes for producers who just fail to qualify for the project given their score. While intuitive and straightforward, this technique only uses observations close to the cut-off point and has not yet featured significantly in any COSA projects.
6. *Sensitivity Analysis* differs for each application but is ideally performed with each of the analytical approaches above to show that the approach was appropriate and to improve the rigor of our results. Furthermore, using an appropriate mix of these techniques when possible (e.g., PSM and DID or PSM and IV) is encouraged.

Stakeholder Consultations

COSA and its local Partner institution conduct a Final Workshop to review the findings with local stakeholders in a focus group format. This provides an opportunity to deepen the understanding of the data and of the local context and to also discover points that may have been missed during field work. Integrating the viewpoints of experts, local people and institutions contributes to new insights, lessons, and a useful validation of the findings. It is also a valuable chance to multiply the benefit of the information gathered by sharing it with those that are directly affected by it. In most cases, COSA data is not final until it is validated by this last step.

Presenting information: Customized KPI and management dashboards

COSA offers real-time management dashboards that are practical and easy for managers to use. They can track standard or customized Key Performance Indicators (KPI) in real time and at the level of detail and nuance that makes sense for them.

Knowledge base & benchmarking

COSA data grows with thousands of new surveys added every year and spans a number of countries. Our Knowledge base houses one of the largest selections of relatively comparable data on agricultural sustainability issues in the world. With forthcoming updates, it will more easily support benchmarking or time-series analyses.

Better informed decisions & policies

COSA's assessments serve to present information in ways that facilitate better informed decisions and policies. We strive to go beyond providing just raw information and seek to distill the critical relationships and important issues with clear and concise determinations of the factors that have the most significant "effect" in a situation.

~ ~ ~