2018

IMPACT AUDIT

TechnoServe

Coffee Initiative Phase II

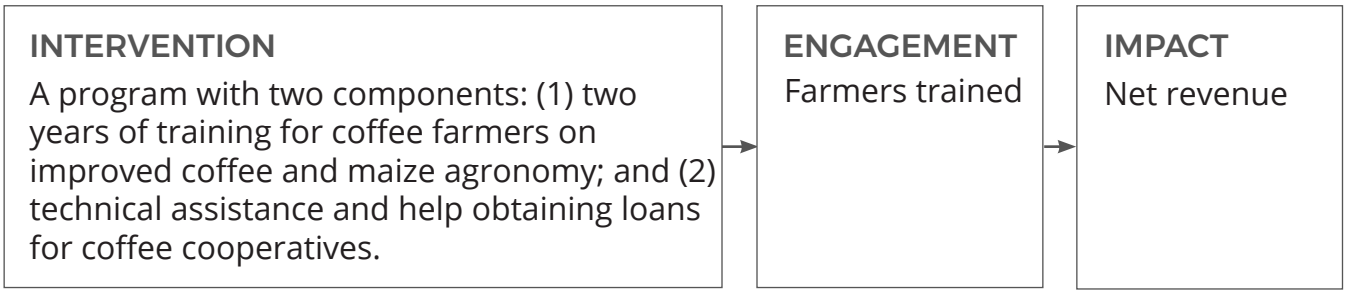
ImpactMatters

FINDINGS

Coffee Initiative Phase II

MISSION To increase the net revenue of smallholder coffee farmers in East Africa.

PROBLEM Smallholder coffee farmers are not aware of agronomic best practices or market opportunities. They also lack access to finance.



IMPACT AND COST

$34 in net revenue per $1 spent

IMPACT AND COST CALCULATION

A 34:1 benefit/cost ratio, calculated from the perspective of TechnoServe, is exceptional. In the calculation, estimated costs include TechnoServe's expenses but not those borne by farmers. When farmers' costs enter the calculation, the benefit/cost ratio falls to 23:1. We assume benefits from the two-year program last for 10 years in total. From the perspective of farmers, net revenue rose $71 for every $1 that farmers paid to participate in the program.

QUALITY OF EVIDENCE

★★☆

QUALITY OF EVIDENCE ASSESSMENT

The evidence for the impact of Coffee Initiative II comes from a medium quality study conducted by TechnoServe. The study surveyed participants and non-participants in Kenya and Rwanda and provided evidence of success of the program. However, participants and non-participants might not have been comparable, since participants volunteered to join the program. Additionally, the results of the study were not applicable to most farmers in the program, who were from Ethiopia. Coffee "stumping," a technique promoted by TechnoServe and a key determinant of yields, was new to Ethiopian farmers, but not to Kenyan and Rwandan farmers.

Table of Contents

1460 Broadway
New York, New York 10036
www.impactm.org • info@impactm.org

Feedback

We welcome your feedback at www.impactm.org/feedback.

EXECUTIVE SUMMARY

Program Description and Key Findings

TechnoServe's stated mission for the Coffee Initiative II (C.I. II) program was "to increase the incomes of smallholder coffee farmers in East Africa." ImpactMatters estimates the mission-driven impact of C.I. II as the increase in the average farmer's net revenues above what he or she would have earned in the absence of the program. The change in net revenue (gross revenues minus the cost of goods sold) is a good proxy for the change in net income (farmer's take-home pay), as explained below.

C.I. II was designed to tackle constraints preventing farmers from maximizing their gains from coffee farming, including their scant financial reserves to invest in vital technologies like wet mills and their lack of knowledge in agronomy and marketing. To address these constraints, TechnoServe provided the following services:

- **Farm College**, a training program that provided two years of monthly sessions on best practices in agronomy, such as weeding, composing and "rejuvenation" (cutting down the tree at its base every eight or so years, allowing it to bear more and better fruit in the long term).
- **Technical assistance, including training and consulting, to farmer-owned cooperatives** on good management practices and how to run wet-mill machinery. Wet mills allow farmers to produce higher quality coffee, which can be sold at a premium.
- **Links to coffee purchasers** interested in buying specialty, wet-milled coffee from cooperatives.
- **Links to financial institutions** willing to lend cooperatives working capital and funds to construct their own wet mills.

This impact audit reviews the second phase of the Coffee Initiative, C.I. II, which was implemented by TechnoServe from 2012 to 2016 in Ethiopia, Kenya, Rwanda and Tanzania. Over that time, C.I. II enabled 185,700 farmers to sell to wet-mill cooperatives, of which about 96,800 also received Farm College training.

ImpactMatters estimates that farmers on average earned an additional $3,300 in net revenues — defined as revenues less the cost of goods sold — over the course of 10 years, above what she or he would have earned in the absence of the program. From the perspective of TechnoServe — therefore counting only those costs incurred by TechnoServe to deliver the program — we estimate a benefit/cost ratio of 34:1, which is exceptionally high. From the perspective of the farmer — therefore counting costs she or he incurs from participating in the program but not the costs covered by TechnoServe — net revenues rose by $71 for every $1 that each farmer spent because they participated in the program (a benefit/cost ratio of 71:1). From the perspective of society-wide impact — therefore counting both the costs incurred by TechnoServe and those incurred by farmers — we estimate a benefit/cost ratio of 23:1.

We judge the evidence behind our estimates to be of medium quality. TechnoServe did not measure net or gross revenues directly. Rather, it collected data on the components of gross revenue: quantity of coffee sold and the price at which it was sold. We therefore model the change in net revenue that arose because of the program as a function of the change in quantity of coffee sold, change in price and the change in costs to farmers of participating in the program. Below, we discuss the quality of the primary piece of evidence in our model: TechnoServe's quasi-experimental study measuring farmers' yields in Kenya and Rwanda.

Impact and Cost

BENEFIT/COST RATIOS

We present three benefit/cost ratios: one from the perspective of TechnoServe, one from the perspective of farmers and one from the combined perspective of all entities that incurred costs because of the program. The numerator is the same in all ratios: $3,300 in additional net revenues per farmer, earned over 10 years. We estimate change in net revenues as a proxy for change in the average farmer's net income. In other words, we assume that if the intervention causes a change in the net revenues of the farm enterprise, the very same change is passed onto by the farmer in her or his net income.

The first ratio takes as its denominator TechnoServe's cost of delivering the program, $100 per farmer. The resulting benefit/cost ratio is 34:1. Every $1 spent by TechnoServe returned $34 in additional net revenues to farmers, net of counterfactual effects (changes in net revenue the farmer would have experienced had she or he, counter to fact, not participated in the program.)

The second benefit/cost ratio takes as its denominator the costs incurred by the average farmer because she or he participated in the program, $50. From the farmer's perspective, the benefit/cost ratio is 71:1. Investing $1 by participating in the program yields a large $71 return for farmers.

Finally, the benefit/cost ratio falls to 23:1, with $140 in the denominator, if we take account of all entities that incurred costs because of the program, meaning both TechnoServe and farmers.

We acknowledge the above benefit/cost ratios contain some uncertainty, both biased upward and downward. The overall direction of the bias is ambiguous without more evidence on how farmers' yields evolve over a longer horizon.

To calculate impact, we draw upon multiple sources of evidence. First, TechnoServe's "difference-in-differences" study that compared the yields of participating farmers and non-participants before and after the intervention. The study found increases in yields of about 67 percent and 41 percent.[i] However, TechnoServe stopped measuring yields immediately after the intervention. To determine the duration of impacts in our model, we turn to the findings of a long-term assessment of farmers' adoption of best practices taught in C.I. II. We infer that yields can be expected to remain higher than they would otherwise be in the absence of the program for as long as more participating farmers than non-participating farmers adopt a certain critical threshold of best practices. In other words, in our model, adoption of best practices determines yields. Finally, we use data collected by program staff on the price premiums farmers reap by selling specialty, wet-milled coffee — between 88 cents and $1.25 more per kilogram of coffee. We reason that farmers continue to enjoy price premiums for as long as cooperatives can be expected to manage and maintain their wet-mill enterprises.

DISPLACEMENT AND OTHER EFFECTS

C.I. II may have triggered other effects not captured in our estimates of the increase in net revenues accruing to farmers.

Besides training farmers in coffee agronomy, C.I. II also trained farmers on nutrition and maize farming, with the intention of bolstering their food security. By the end of the program, fewer participating farmers than non-participants experienced food shortages and many participants were observed reinvesting their additional coffee revenues into maize, which they both consumed directly and traded for other foodstuffs.

[i] Significant at a confidence level of 95 percent and presented as a percentage of the comparison group's average yield after the program.

C.I. II also trained farmers in environmentally friendly farming techniques, such as mulching, composting and prudent use of organic fertilizers. These techniques might have improved biodiversity and the quality of soil and water. And like many TechnoServe programs, C.I. II had an explicit focus on gender equality, seeking out women farmers to not only receive training, but also take leadership roles in the farming community.

We do not think C.I. II farmers displaced other farmers selling specialty coffee. The global market for specialty coffee is growing. Trade has been increasing for the more expensive, sought-after arabica coffee species and decreasing for the cheaper, commodity-grade robusta species. C.I. II farmers entering the specialty coffee market are likely fulfilling ample demand rather than displacing other farmers.

Quality of Evidence 

The primary source of evidence that undergirds our estimates of impact is TechnoServe's own differences-in-differences study of C.I. II. TechnoServe surveyed participating and non-participating farmers in Kenya and Rwanda and found that yields of participating farmers increased substantially more than did those of non-participants after receiving the program.

We downgrade the quality of this study for two key reasons.

First, we suspect selection bias caused the survey results to overstate increases in yields during the two years of observation. Farmers self-selected to receive training rather than being randomly assigned to a treatment and control group. Farmers who volunteered may be different from non-volunteers in important ways that determine yields and net revenues, such as their motivation and resourcefulness.

Second, the farmers surveyed in Kenya and Rwanda were not representative of the total population of farmers who participated in C.I. II, three-quarters of whom were based in Ethiopia. In contrast to Kenyan and Rwandan farmers, farmers in Ethiopia were new to the best practice of "stumping"— cutting the tree at its base and leaving it out of production for about two years before it starts bear more and better fruit in future years. The extrapolation of results from Kenya and Rwanda to Ethiopia is imperfect, likely predicting too high an increase in yields during the years when stumped trees do not bear fruit and predicting too low an increase in yields when stumped trees begin to flourish. We are unable to quantify the magnitude of these biases and the extent to which one might offset the other.

NONPROFIT COMMENT

[PLACEHOLDER FOR COMMENT FROM TECHNOSERVE ON THE REPORT]

PROGRAM DESCRIPTION

This section summarizes the program's mission and constructs a theory of change that describes the problem, TechnoServe's intervention and ImpactMatters' chosen measure of impact.

Mission

To increase the net revenue of smallholder coffee farmers in East Africa.

TechnoServe's stated mission of Coffee Initiative II is "to increase the incomes of smallholder coffee farmers in East Africa."[1] ImpactMatters measures achievement of this mission as the increase in net revenues caused by the program for participating farmers, as explained below.

Theory of Change

PROBLEM

Smallholder coffee farmers in East Africa are not aware of best practices in agronomy or of the opportunities available to them in the market for high quality coffee. They also lack access to finance to invest in vital technologies like wet mills, which are used to separate the skin and pulp from the coffee fruit. Wet mills yield more consistent, higher quality coffee than sun-drying or partial washing.

LACK OF AGRONOMY KNOWLEDGE

The primary reason for poor productivity in the East African coffee sector is farmers' low adoption of best practices in agronomy.[2] In TechnoServe's experience, farming communities have never had access to formal training on even the most basic of farming

skills. Instead, farmers rely on generations-old farming practices and what they observe from their neighbors.

LACK OF MARKET KNOWLEDGE

Farmers' knowledge of the market for coffee is limited to their immediate line of sight: growing commodity-grade coffee to be sold to an intermediary. They are unaware that there is a market for higher quality coffee, especially wet-milled coffee and coffee that has been "cupped" by a professional to evaluate its taste and aroma.[3] As a result, farmers unknowingly forgo the price premiums they would obtain from selling specialty coffee.

LACK OF ACCESS TO FINANCE

Even if farmers are fully aware of the market opportunities they could enjoy by setting up wet mills and cupping labs, they do not have the funds for such capital investments.[3] Farmers and cooperatives are not well-positioned to negotiate for loans from banks, especially as they do not have loan guarantors or collateral to offer.[4]

ACTIVITIES

During Coffee Initiative Phase II (C.I. II), coffee farmers attended "Farm College," two years of training on improved coffee and maize agronomy. Simultaneously, the program provided "technical assistance" (training and consultancy) to farmer-owned coffee cooperatives, including introducing wet-mill technology. The program also helped cooperatives obtain loans and reach purchase agreements with specialty coffee purchasers.

FARM COLLEGE

Farm College consisted of two years of monthly training sessions, delivered by trainers employed by TechnoServe.[2] Training was delivered to groups of about 30 farmers and took place on the coffee field of the "focal farmer," a respected coffee farmer selected by each group of 30.

Farmers were trained in improved coffee agronomy practices like weeding, composting and "rejuvenation," cutting the trunks of the coffee tree to allow it to bear more and better fruit about two years later. Training also included four modules on maize farming and one module on infant nutrition.

TECHNICAL ASSISTANCE TO COOPERATIVES

Business advisors employed by TechnoServe provided technical assistance to cooperatives, including support on: wet-mill technology and management; cooperative management and governance; quality control of coffee; and social and environmental responsibility.[2]

Business advisors also advised cooperative unions, which collect coffee from cooperatives to sell to purchasers, on how to improve management practices, compliance with regulations and marketing.[5]

LINKAGES TO COFFEE PURCHASERS AND FINANCIAL INSTITUTIONS

TechnoServe promoted coffee produced by participating cooperatives to buyers of specialty coffee. It helped cooperatives find loan guarantors and negotiated with local banks so that cooperatives could borrow money for both day-to-day operations and investment in wet mills.[2]

MEASURES OF IMPACT

In this impact audit, the success of C.I. II is measured as the **increase in farmers' net revenue from coffee.** This metric captures both (1) the increase in farmers' *yields* due to improved agricultural techniques and (2) the increase in *price*, known as the "price premium," that farmers receive for selling higher quality coffee.

Net revenue is defined as *value of coffee sold* (gross revenue) minus the *cost of producing coffee* (commonly termed "cost of goods sold"). Both TechnoServe and academic researchers report difficulty measuring the latter, the cost of production for smallholder farmers.[6] Our approach circumvents this difficulty by focusing on rates of change. We need only estimate the *change* in gross revenues caused by the program minus the *change* in all other costs caused by the program to arrive at the *change* in net revenues caused by the program. We then treat the change in net revenues as a proxy for change in the average farmer's net income. In other words, we assume that if the intervention causes a change in the net revenues of the farm enterprise, the very same change is passed onto by the farmer in his net income.

We do not have enough data to accurately determine how much farmers' costs rose because of the program. Our calculation of the change in net revenue might, therefore, be incomplete. While we have information on gross revenues, we make some assumptions about how farmers' costs changed from participating in the program. TechnoServe staff and documents indicate C.I. II promoted virtually costless best practices, so we infer that

farmers' cost of goods sold simply did not change because of participation.[3,7] However, we know C.I. II caused some farmers in Ethiopia to join coffee cooperatives, for which they paid a one-time joining fee. Farmers also gave up time to attend training. We assign monetary value to both of these costs, described in the section on Impact and Cost.

TechnoServe's mission as an organization is to create more competitive farms, businesses and industries. It aims to increase economic activity across market systems, not limited to the market actors that participate directly in its programs. TechnoServe measures system-wide effects as the change in gross revenues of its program participants, reasoning that a share of those revenues is distributed to the farm laborers, financial institutions and other market actors with whom participants do business. As such, in TechnoServe's view, our chosen measure of impact (net revenues accruing to program participants) does not capture the full extent of its intended market effects.

We do not agree that change in gross revenue is a useful indicator, whether of benefits to the participant or to the market system. Focusing on change in gross revenue overlooks change in costs caused by the program. Further, the research community has yet to confirm the system-wide benefits of such programs as TechnoServe's. We believe our analysis of the change in net revenues accruing to program participants is an appropriate reflection of the impact of TechnoServe's programs. In the section on Displacement and Other Effects (in the Impact and Cost chapter), we also briefly discuss the potential benefits (and harms) to third parties.

ASSUMPTIONS

Several assumptions had to be true in order for the programmatic activities of C.I. II to produce the participant engagement and impact stated above.

TechnoServe assumed that farmer trainers and business advisors provided accurate information and delivered services professionally. It also assumed that buyers of specialty coffee would be willing to buy from small farmers with no track record of success. Similarly, it assumed banks would be willing to provide loans to small farmers who may not have had a credit history and who were working in the highly volatile coffee sector. Loan guarantors were also assumed to be willing to underwrite loans for small farmers.

RISKS

The biggest risk that threatened the impact of C.I. II was the risk of environmental shocks such as flood, drought and crop disease. Though TechnoServe could not eliminate these risks entirely, it did train farmers on certain "climate-smart" best practices. For instance, when temperatures surged, farmers used mulch and shade to maintain the right temperatures for their soil and plants.[2]

Environmental shocks elsewhere in the world could also have affected C.I. II farmers. A bout of good weather can temporarily boost the yields of farmers elsewhere in the world. When those yields flood the market, global prices of coffee plummet.

Program Details

GEOGRAPHY

C.I. II is the second phase of a two-phase program. The two phases are identical, except that C.I. II also included some training on maize agronomy on top of the other program activities.

Phase one took place in Ethiopia, Kenya, Rwanda and Tanzania. In C.I. II, all the activities of phase one were continued in Ethiopia. All programmatic activities in Kenya, Rwanda and Tanzania were wound down over a span of two years. Ethiopia was chosen for C.I. II because of the preferences of the program's donor. Similarly, training in the farming of maize was introduced in C.I. II because of the donor's interest in food security. In East Africa, almost all coffee farmers also grow staple food crops like maize.[2] Unlike coffee, maize is not a cash crop, but a source of food consumed by farmers themselves. TechnoServe found farmers were investing their extra coffee income, now higher after training, into their maize farms. To maximize that reinvestment and shore up food security, TechnoServe introduced maize training.

STAGE

C.I. II reached the "scale" stage.[i] Phase one of the program underwent two major independent assessments to validate its success. Equipped with the largely positive results from those assessments, TechnoServe decided to expand the program with C.I. II.

AGE AND SCALE

C.I. II enabled 185,700 farmers to sell to wet-mill cooperatives, of which about 96,800 also received Farm College training.

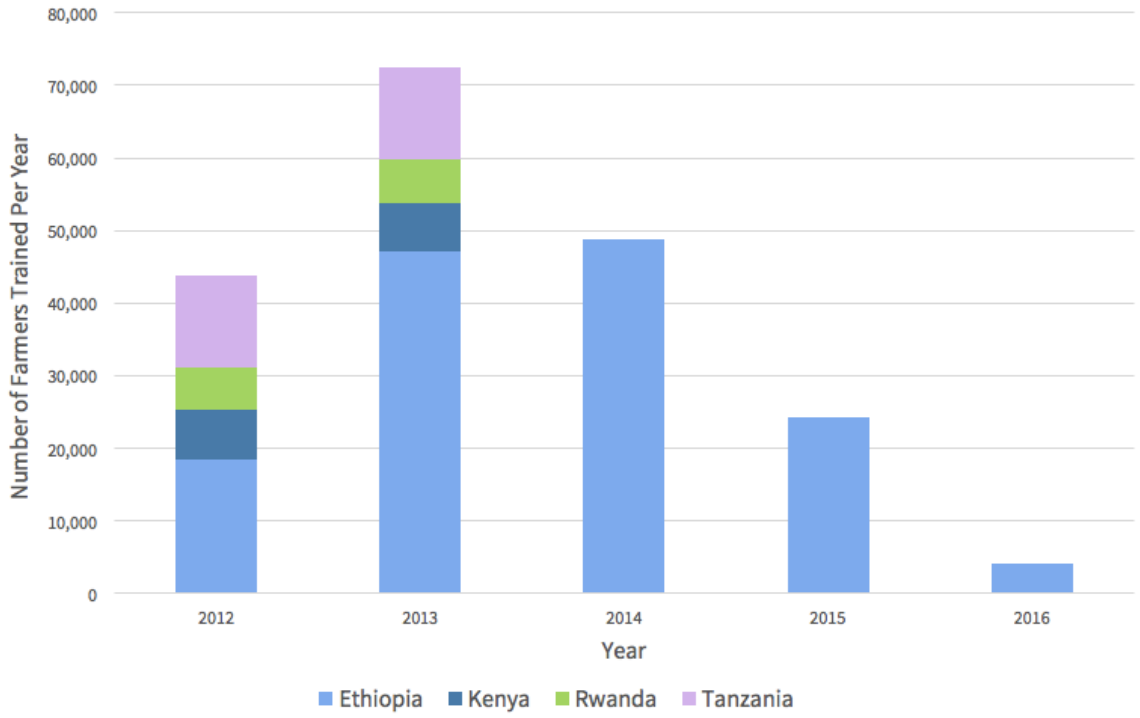


Figure 1. Number of farmers trained, 2012-16

One cohort each in Kenya, Rwanda and Tanzania received the intervention in 2012 and 2013, after which program activities in those countries ended.

The intervention was delivered to overlapping cohorts of farmers in Ethiopia, with each cohort participating in the program for two years. Cohorts varied in size. A new cohort

[i] ImpactMatters classifies programs on a continuum from "design" stage to "validation" and "scale." At the design stage, the nonprofit is discovering the right way to implement its intervention. A program at the "validation" stage is being tested for cost-effectiveness, before the nonprofit expands access to the program. At the "scale" stage, the nonprofit is expanding access to the program, to the extent warranted by its cost-effectiveness.

entered the program every February from 2012 to 2015, such that over the program's lifetime, C.I. II served four cohorts in Ethiopia.

FUNDING

TechnoServe spent $18 million to deliver C.I. II, or an annual average of $3.6 million over its five years. By our calculations, C.I. II accounted for about 4 percent of TechnoServe's total annual expenses over that time period.

IMPACT AND COST

WHY WE ESTIMATE

Impact audits estimate the philanthropic impact and cost of a nonprofit's programmatic interventions. We base those estimates on best available evidence, however imperfect, drawn from the auditee (internal evidence) and research literature (external evidence). As such, our estimates are the best possible evidence-based gauge of philanthropic success.

HOW WE ESTIMATE

First, we identify outcomes that best capture the auditee's mission. We then settle upon ways to measure progress against those outcomes, relying on the tools of modern social science.

Second, we report our estimate of "impact," the change in outcomes that can be attributed to the auditee's intervention over a designated period of time. We take explicit account of counterfactual success — the change in outcomes that would have occurred without the program. And whenever possible, we take explicit account of third-party effects, especially unintended harm to vulnerable individuals because of the auditee's intervention. For benefits that accrue over time — for example, the increased earnings from high school graduation — we discount these future benefits (at a 5 percent discount rate). The length of time over which benefits are assumed to accrue is based on the specifics of the intervention under review and available internal and external data.

Third, we report total costs. Total costs include marginal costs (direct costs of delivering the intervention) and fixed costs (for example, administrative overhead) regardless of who bears those costs (nonprofit, public agencies, private funders or participants). For programs that generate commercial revenue, the revenue is treated as a subtraction of costs. For costs that kick in over time, we discount (as we do benefits). The length of time over which costs accrue depends on the specifics of the intervention under review and available internal and external data.

Fourth, we report the ratio of impact to cost (a benefit/cost ratio).

Finally, we analyze key factors — for example, stage of development, whether the nonprofit be in pilot phase or expansion phase — relevant for understanding the audit findings.

Typically, impact is estimated on a single outcome. However, if an auditee's intervention affects several outcomes, we report impacts on distinct outcomes separately. Concretely, suppose that a program seeks to raise incomes and improve health status. We do not, as yet, attempt to combine the impact on multiple outcomes into a single aggregate outcome — concretely, by combining the value of the income effects and health-status effects. To aggregate, we would need weights — the relative value of outcomes — that would reflect the nonprofit's or funder's values (not those of ImpactMatters as auditor).

Findings

ImpactMatters measures the impact of Coffee Initiative II (C.I. II) as the increase in net revenue of participating coffee farmers in East Africa, net of the gains or losses they would have experienced had they not in fact participated in the program. The stated mission of TechnoServe's Coffee Initiative II (C.I. II) program is to "increase the incomes" of smallholder coffee farmers in East Africa. We calculate change in net revenues as a next-best alternative that we think closely tracks movement in income.

We estimate net revenue per farmer increased by $34 for every $1 that TechnoServe spent on the program (a benefit/cost ratio of 34:1).[i] We estimate that C.I. II increased net revenues for the average farmer by $3,300 over the course of 10 years (two years spent in the program and eight years thereafter). A benefit/cost ratio of 34:1 is exceptionally high. In the Quality of Evidence section of this report, we discuss reasons why this estimate carries some uncertainty, though we cannot be sure whether the estimate is biased upwards or downwards.

The benefit/cost ratio can also be calculated from the perspective of the farmer. We estimate net revenue per farmer increased by $71 for every $1 paid by the farmer (a benefit/cost ratio of 71:1). If we take account of all costs incurred because of the program,

[i] All figures are presented in 2016 U.S. dollars. Figures originally denominated in Ethiopian birr, Kenyan shillings, Rwandan francs and Tanzanian shillings were converted at purchasing power parity.

both those incurred by TechnoServe and those incurred by farmers, the benefit/cost ratio falls to 23:1.

Table 1. Impact and Cost Findings

Specification	Total
TOTAL FARMERS REACHED DURING PROGRAM	185,700
INCREASE IN NET REVENUE (OVER 10 YEARS)	$3,300
BENEFIT/COST RATIO, COUNTING ONLY COSTS COVERED BY TECHNOSERVE	34:1
BENEFIT/COST RATIO, COUNTING ONLY COSTS COVERED BY FARMERS	71:1
BENEFIT/COST RATIO, COUNTING ALL COSTS, REGARDLESS OF PAYER	23:1

We define the numerator of the benefit/cost ratio as the change in net revenue due to the intervention: the change in the value of sales minus the change in costs due to the intervention. For a coffee farmer in the program, that means the change in the value of coffee sold (over the last year) minus the change in the cost of seeds, fertilizer, farm equipment, farm labor and the like, that the farmer incurred to produce the coffee he sold.

Staff testimony[3] and documents from TechnoServe[7] say that farmers spent negligible additional money to apply the best practices taught in C.I. II. ImpactMatters assumes this is true and that the only additional costs to farmers as a result of their participation in the program are: (1) labor-market income forgone from the time they spent in training; and (2) the fees they paid to join coffee cooperatives. We estimate, based on TechnoServe's data and staff interviews, that about three-quarters of farmers in Ethiopia joined cooperatives because of TechnoServe's intervention.[7]

As stated in the table above, we estimate that the intervention increased net revenue by $3,300 per farmer over 10 years, meaning the growth in sales revenue surpassed the change in costs by $3,300 over seven years.

We estimate three versions of the benefit/cost ratio. One is from the perspective of TechnoServe; the second is from the perspective of the typical farmer; and the third is from a societal perspective, taking account of all costs incurred because of the program,

regardless of who bears the costs. In all versions, the above-mentioned $3,300 figure serves as numerator.

The first version, calculated from TechnoServe's point of view, takes as its denominator the amount that TechnoServe spent to implement the program divided by the number of farmers who participated. That cost amounts to $100 per farmer. The second version, calculated from the perspective of the farmers, takes as its denominator the $50 paid by each farmer in fees and forgone earnings. Finally, the third benefit/cost ratio takes as its denominator $140, which combines TechnoServe's costs and farmers' costs incurred because of the program.

STRATEGY FOR ESTIMATING IMPACT

C.I. II raised the net revenues of participating farmers in two ways. First, Farm College training increased farmers' adoption of best practices in agronomy, which led directly to an increase in **coffee yields** (amount of coffee produced per area of land). Second, with TechnoServe's assistance, farmer-owned coffee cooperatives were able to set up wet mills. Cooperatives charged buyers a **premium price** for their wet-milled coffee, a specialty item compared to regular sun-dried coffee. Part of that price premium was passed on to the farmers who sold their coffee to the cooperatives.

About half of all farmers received the two components of the intervention, Farm College and access to a wet mill.[8] The other half sold to wet-mill cooperatives but did not attend Farm College.

Impact per participating farmer must therefore be disaggregated into two models:

For farmers who both attended Farm College and sold to wet-mill cooperatives:

$$I_{Farm\ College\ +\ wet\ mill} = [Y * F * P_W] + [V * (P_W - P_M)] - C$$

For farmers who did not attend Farm College, but did sell to wet-mill cooperatives:

$$I_{wet\ mill\ only} = [V * (P_W - P_M)] - C$$

where:

I = Impact measured as the increase in net revenue per farmer (dollars)

Y = Increase in yield per farmer over one year (kg per tree)

F = Average productive farm size per farmer (trees)

P_W = Price of wet-milled coffee (dollars per kg)

V = Average amount of coffee sold per farmer over one year, measured before the program (kg)

P_M = Market price of sun-dried coffee (dollars per kg)

C = Change in cost of sales per farmer from participating in the program (dollars), which we assume only includes the farmer's opportunity cost of time spent in training and fees to join a coffee cooperative

In the first model, the boost to revenues is driven by both higher yields, Y, and higher prices obtained for wet-milled coffee, or the difference between P_W and P_M. Farmers in the second model rely solely on higher prices to earn higher revenues.

The annual impact of C.I. II on the average participant is the weighted average of the two models:

$$I = \frac{N_{FC+WM} * \left[(Y * F * P_W) + \ (V * (P_W - P_M))\right) - C] + N_{WM} * [(V * (P_W - P_M)) - C]}{N_{FC+WM} + N_{WM}}$$

where:

N_{FC+WM} = Number of farmers who both attended Farm College and sold to wet-mill cooperatives

N_{WM} = Number of farmers who did not attend Farm College but sold to wet-mill cooperatives

C.I. II did not measure the net revenues of farmers directly, such as by asking farmers to self-report how much they made by selling coffee in the last year. TechnoServe staff believed such an approach would result in extremely inaccurate estimates of impact because farmers tended not to keep any records of sales made.[3] Instead, TechnoServe conducted a "quasi-experimental" study of how farmers' yields changed in response to the intervention.[9,10] This study serves as one of the key ingredients in our benefit/cost calculation. It directly informs the value of the variable Y (increase in yield) in the algebraic model above.

In the quasi-experimental study, TechnoServe surveyed a group of C.I. II farmers and a group of comparable farmers before and after the program. The surveys took place in Kenya and Rwanda, two of the four countries where C.I. II was implemented. TechnoServe used a "difference-in-differences" method to compare the two groups across time. The results: an impressive 67 percent and 41 percent increase in yields in Kenya and Rwanda,

respectively.[i] The study lacked random assignment of farmers into treatment and control groups, making it a quasi-experimental study, in contrast to a randomized controlled trial. We evaluate the quality of this study in the Quality of Evidence section below.

The next key input to our analysis is the lift in price that C.I. II farmers obtain from selling wet-milled coffee instead of sun-dried coffee. TechnoServe field staff observed the prices that farmers obtained for the two different types of coffee. They found wet-milled coffee earned a price premium ranging from 88 cents to $1.25 above the price of sun-dried coffee.[11][ii]

Finally, we model the effects of C.I. II over time: After two years of Farm College training and support to set up a wet mill, for how long do productivity and financial benefits persist for farmers? We assume that the yields of non-participants and participants converge 10 years after the start of the program. In other words, impacts on yields cease after 10 years. To make our assumption, we proceed in four steps:

1. At the seven-year mark, independent evaluator Triple Line visited randomly selected farmers who had participated in the first phase of the Coffee Initiative in Rwanda.[11] Triple Line observed continued adoption of best practices in agronomy. By our calculations, which are based on Triple Line's data, an average of 23 percent of farmers adopted the critical threshold of at least eight out of 10[iii] best practices at the seven-year mark,[iv] compared to 55 percent immediately after the intervention and 7 percent before the intervention. We assume a linear decrease in adoption from 55 percent (immediately after the intervention) to 23 percent (at the seven-year mark), meaning adoption fell in equal amounts from one year to the next. We extend the trend-line to future years, illustrated by the blue curve in

[i] Significant at a confidence level of 95 percent and presented as a percentage of the comparisonl group's average yield after the program.

[ii] We use the weighted average of coffee prices in Rwanda and Ethiopia, weighted by the number of participating farmers in each country. The (simple) average of prices in Rwanda and Ethiopia serve as the prices in Kenya and Tanzania, for which price data were not collected.

[iii] In their assessment of the first phase of the Coffee Initiative in Rwanda, independent research consultants Laterite Ltd.[18] found farmers who adopted more than seven (out of a total of 11) best practices tended to have similar improvements in yields. We have insufficient information to quantify the effects on yields of each additional best practice adopted (and adopted in what combination). But we can roughly assume, based on Laterite Ltd.'s findings, that a minimum threshold of eight best practices adopted leads to higher yields.

[iv] There were 11 best practices in total, but Triple Line was only able to observe 10. Triple Line reported the mean adoption rate across all best practices observed. We used the reported mean to calculate the percentage of farmers who adopted eight or more best practices, assuming a Poisson distribution consistent with TechnoServe's evidence from Ethiopia.[14]

Figure 2 below.

2. Triple Line did not survey a comparison group. Its evaluation does not demonstrate effects relative to a counterfactual case — farmers' rate of best practice adoption if they had not participated in the program. We therefore draw from TechnoServe's own surveys of best practice adoption in Ethiopia, Kenya, Rwanda and Tanzania, which did include a comparison group but tracked farmers for only the two years of the intervention. The surveys show the percentage of comparison-group farmers adopting the critical number of best practices increased from 0.9 percent to 1.3 percent over the period of intervention.[i] We assume this is the constant rate of "catch up" — the natural rate at which farmers who did not receive the intervention adopt best practices on their own. The catch-up of the comparison group is illustrated by the gray curve in figure 2 below.

3. The linear decline of the treatment group meets the gradual catch-up of the comparison group in between years 10 and 11. Said differently, 10 years after the start of the program, the percentage of farmers adopting the critical threshold of best practices is equal in the treatment and comparison groups. This is illustrated by the convergence of the blue and gray curves in figure 2 below.

4. We assume the trajectory of yields mirrors the trajectory of adoption of the critical number of best practices. The impact of C.I. II on farmers' yields therefore also ends after 10 years. Figure 3 illustrates this trajectory.

[i] Average weighted by the number of farmers participating in each of the four countries.

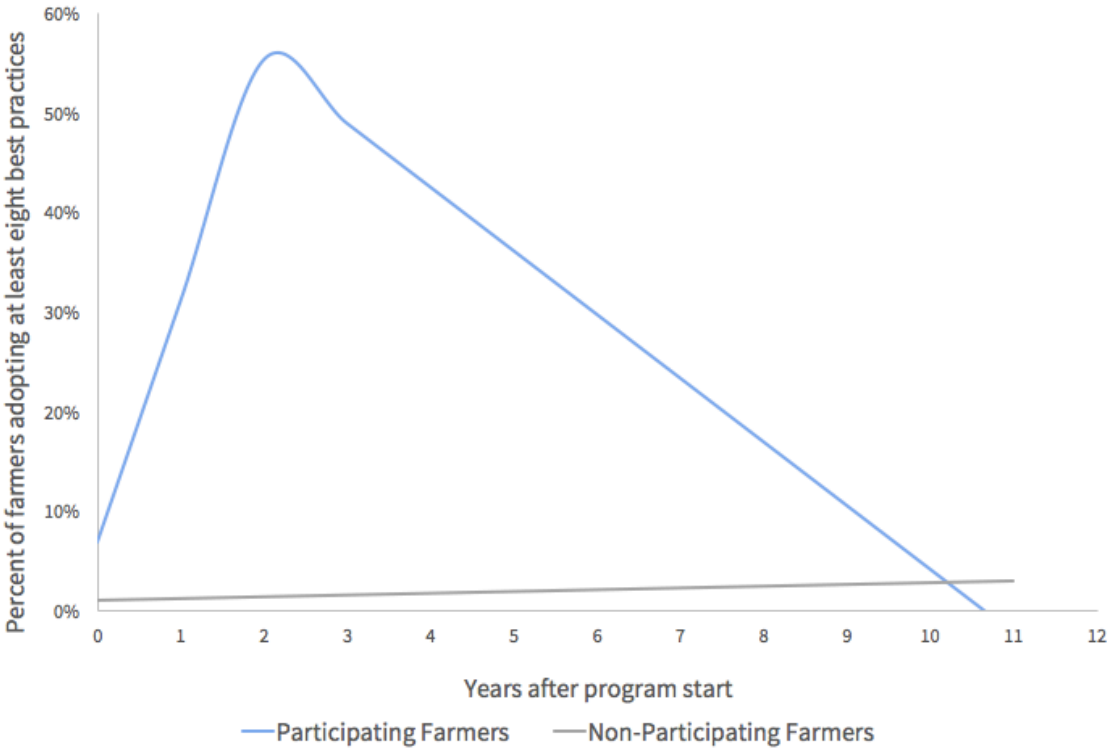


Figure 2. Best Practice Adoption Among Participating and Non-Participating Farmers

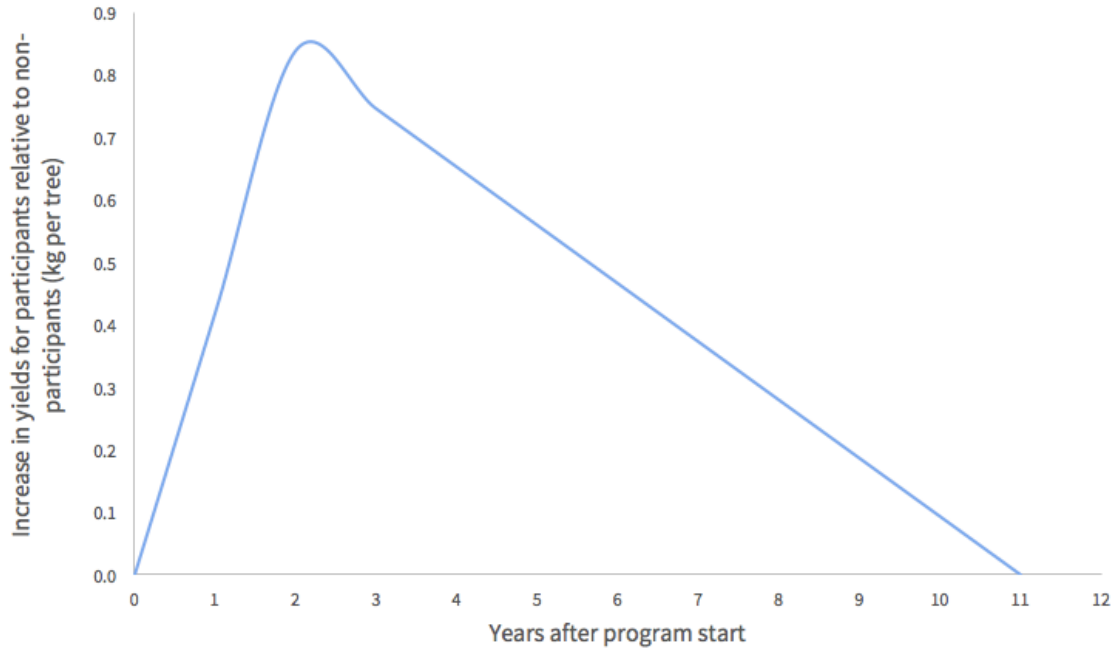


Figure 3. Increase in Yields of Participating Farmers Relative to Non-Participants

Recall there are two components of C.I. II: Farm College and access to wet mills. We estimate that impacts driven by access to wet mills continue for up to eight years after the start of C.I. II. Unlike agronomy training, which relies on farmers to continuously apply their newly acquired knowledge of best practices, access to a wet mill can be considered a one-time intervention. Once the wet mill is installed at the cooperative, farmers wash their coffee at no extra cost and can sell their wet-milled coffee at a premium. The longevity of impact driven by the wet-mill component of C.I. II depends on three factors: (1) whether non-participating farmers gain access to wet mills, independent of the program; (2) the durability and maintenance of the wet mills; and (3) how well the cooperatives that run the wet mills manage themselves.

1. According to interviews with program staff[3] and the Triple Line evaluation,[11] C.I. II was largely responsible for the introduction of wet-mill technology to the cooperatives targeted in Ethiopia, where almost three-quarters of all farmers trained by the program were from. Had TechnoServe not intervened, those cooperatives would likely not have set up wet mills. We therefore assume that other farmers do not gain access to wet mills in the near term.
2. At the seven-year mark, Triple Line visited a non-random sample of 10 wet-mill cooperatives (five each in Ethiopia and Rwanda). Four in Rwanda and all five in Ethiopia had wet mills that appeared to be in good condition and were being well maintained.[11] However, Triple Line noted that cooperatives in Rwanda tended to buy new coffee pulpers instead of regularly servicing and maintaining the wet mills. They also reported poor availability of spare parts in the region, leading cooperatives to unnecessarily purchase entire pulpers anew. One cooperative in Rwanda also had one idle wet mill.
3. Triple Line rated the 10 cooperatives on "sustainability" — whether the cooperatives needed further management and technical support to handle day-to-day operations. Of the 10, three in Ethiopia and zero in Rwanda were considered "fully sustainable." Two cooperatives each in Ethiopia and Rwanda were deemed to be "doing satisfactorily, but [in] need [of] some support." Three in Rwanda needed substantial support. Nevertheless, profit and loss data from all 89 wet-mill cooperatives supported in Ethiopia showed cooperatives remained profitable over five years of data collection (2009-14), with a gradually increasing trend.

We consider the evidence encouraging, but mixed. The biggest risk to the longevity of impact seems to be poor cooperative management and maintenance of machinery. But without a random sample of cooperatives, we are unsure how representative the Triple Line results are of all the cooperatives in the program. Faced with these results, we

assume that impacts driven by wet mills last for eight years, which is one more year after Triple Line observed nine out of 10 cooperatives had wet mills in good condition.

CALCULATIONS

We calculate impact separately for the two groups of farmers: those who both attended Farm College and sold wet-milled coffee, and those who did not attend Farm College, but did sell wet-milled coffee.

Farmers who attended Farm College received a boost in yields of 0.84 kilograms per coffee tree after the intervention, using the weighted average results of TechnoServe's quasi-experimental study. The average farmer had about 240 trees, a quarter of which we assume were out of commission due to the "stumping" practice taught at Farm College, whereby trees are cut close to the base so that they produce even greater yields about two years later. Our calculations show farmers who attended Farm College produced 150 more kilograms of coffee than non-attendees after TechnoServe's intervention. Three years later, they produced 100 kilograms more than non-attendees; three more years later, they produced 50 kilograms more than non-attendees; and in year 11, the yields of participants and non-participants converged.

At the same time, all the farmers who attended Farm College also sold their coffee to wet-mill cooperatives. We multiply the amount of coffee they would have sold *without the boost in yields* against the price premium to figure the impact driven by the price premium. As explained above, farmers benefit from the price premium for a total of eight years.

The average farmer who both attended Farm College and sold wet-milled coffee earned about $5,300 more in net revenue than a farmer who did not participate in C.I. II, over the course of 10 years. That figure is calculated by subtracting the increase in the farmer's costs, $70, from the increase in his topline sales revenue, $5,400. We estimated the $70 cost based on TechnoServe's data and interviews with program staff, which indicated some 75 percent of farmers in Ethiopia joined cooperatives as a result of participating in the program and paid one-time joining fees of about $50 each.[7,8] In addition, female farmers spent time usually reserved for household chores participating in agronomy training instead. We assume their domestic work freed up other members of the household to engage in income-generating activities, so we apply the appropriate national minimum wage to those hours forgone.

The average farmer who did not attend Farm College but did sell wet-milled coffee earned $1,100 more in net revenue than a non-participant, over the course of 10 years. We

calculated that figure by subtracting the average increase in costs per farmer, $30, from the average increase in topline sales revenue per farmer, $1,200 (note: figures rounded).

The almost fivefold difference ($5,300 versus $1,100) is expected, since most of the impact of C.I. II is driven by dramatic increases in yields found in the quasi-experimental study TechnoServe conducted rather than the relatively smaller increment in price per kilogram for specialty coffee.

In total, 96,700 farmers benefitted from the $5,300 increase in net revenue, while 89,000 benefitted from the $1,100 increase. The weighted average increase in net revenue per C.I. II farmer is $3,300, earned over 10 years. This is the numerator in all three versions of our benefit/cost ratio.

The denominators of the three versions are different. In the first benefit/cost ratio, the denominator is the average cost to TechnoServe of supporting one C.I. II farmer. That cost was just $100, resulting in a high benefit/cost ratio of 34:1. The second benefit/cost ratio is calculated from the farmer's point of view. The average farmer paid $50 in fees and forgone labor-market income to increase his net revenue, resulting in an even higher benefit/cost ratio of 71:1. Adding together TechnoServe's costs and farmers' costs gives a denominator of $140 and a benefit/cost ratio of 23:1. No other entities besides TechnoServe and farmers incurred costs in pursuit of impact.

At $18 million in total costs to TechnoServe[i] and $8.7 million in total costs to farmers, C.I. II was by no means a small program. But because it reached an impressive 185,700 farmers, it was able to achieve a low cost per farmer. Moreover, almost half of those farmers were "passive" participants in C.I. II who did not attend Farm College trainings, but who benefitted from wet mills installed at their cooperatives. Note that TechnoServe did not fund the wet mills directly, but simply trained cooperatives on wet-mill technology and management, and helped them find loan guarantors and banks willing to fund wet mills — a far less expensive intervention on TechnoServe's part. We suspect the wet-mill component is the most cost-efficient component of C.I. II, but are unable to draw a conclusion without being able to break out the programmatic cost of wet-mill activities from Farm College activities.

[i] Our figures are presented in 2016 U.S. dollars and may therefore appear slightly different from TechnoServe's financials.

Displacement and Other Effects

DISPLACEMENT

EFFECT: NONE

We have no reason to suspect that C.I. II farmers displaced other coffee farmers selling specialty coffee. The global market for specialty coffee is growing, with trade increasing by 4 percent from 2015 to 2016 for the more expensive, sought-after arabica coffee species, compared to the 8 percent decrease in trade of the cheaper, commodity-grade robusta species.[12] C.I. II farmers entering the specialty coffee market are likely fulfilling ample demand rather than displacing other farmers.

FOOD SECURITY

EFFECT: POSITIVE

C.I. II also included some training on nutrition and maize farming, with the intention of bolstering the food security of coffee farmers participating in the program. Participants would be able to consume the maize directly and trade any surplus maize for other foodstuffs. Furthermore, nearly 87 percent of C.I. II farmers supplemented their family's food supply by using their coffee income to purchase additional maize.[13] 83 percent of farmers also reported that training in coffee agronomy resulted in the increased yields of other crops. Indeed, one year after training, about 40 percent of participating farmers reported food shortages during the heavy summer rains, compared to 60 percent of non-participants.[14]

ENVIRONMENTAL BENEFITS

EFFECT: POSITIVE

C.I. II trained cooperatives to use wet mills with systems for wastewater treatment. The water released from such wet mills improves the quality of surface water, in contrast to untreated water released from traditional home-processing methods.[15] The program also encouraged best practices like mulching, composting, erosion control and prudent use of organic fertilizers, which were beneficial to both surface-water availability and soil health.

C.I. II helped restore species diversity by encouraging farmers to plant indigenous tree species. In addition, during the one or two years when recently stumped coffee trees (trees cut close to the base to promote greater yields in future years) were not producing coffee cherries, C.I. II trained farmers on "intercropping."[8] Farmers took advantage of the extra light that could now reach the soil around stumped trees and planted other crops

such as beans and chilies. This not only improved biodiversity and soil quality, but also gave farmers an extra source of food and revenue while waiting for stumped trees to grow back.

GENDER EQUALITY

EFFECT: POSITIVE

Women in East Africa have less access to agricultural training, farm inputs and credit than men do.[2] Cultural norms also often label coffee farming as "men's work," preventing women from participating in decisions about the household's coffee trees and sharing in coffee revenues. To bring greater gender equality, C.I. II deliberately sought out women trainees and placed them in leadership roles in the community.

C.I. II aimed to train farmers at a ratio of 70 percent men and 30 percent women. It exceeded this goal by the end of the program, with women making up 42 percent of trained farmers in Ethiopia.[4] Throughout the program, men were encouraged to bring their wives to each training session.[2] And in each training group of 30 farmers, a woman was elected to serve as either the "focal farmer" or "deputy focal farmer," lead farmers who would host the training sessions at their farms.

QUALITY OF EVIDENCE

WHY WE RATE

Quality of evidence reflects our confidence in the impact and cost estimates. For programs with high quality evidence, the impact and cost estimates are more likely to accurately reflect the effectiveness of the program. Quality of evidence reflects only that data we used to construct the impact and cost estimate.

HOW WE RATE

Quality of evidence is rated using an adaptation of the GRADE methodology, a systematic approach to judging evidence. Initially, studies are ranked by whether they are observational, quasi-experimental or experimental. Then, each study is assessed against quality criteria: risk of bias, inconsistency or results, indirectness of evidence, imprecision, risk of publication bias, magnitude of effect, evidence of a dose-response relationship and attenuation bias.

In the ideal case, data from the program are solely used to estimate the impact of the program. However, external data can be used to identify quantitative and qualitative parameters or to link behavior change to outcomes. When the analysis is substantively based on data from multiple sources, the quality of each is assessed. If only very-low-quality internal data is available, high-quality external data may be substituted. In addition, external evidence can serve to confirm or contradict internal evidence.

Star Rating	Quality of Evidence
☆☆☆	Quality rating is “very low”; or
	Quality rating is “low” but high quality external evidence contradicts its findings
★☆☆	Quality rating is “low”; or
	Quality rating is “medium” but high quality external evidence contradicts its findings; or
	Quality rating is “very low” but high quality external evidence corroborates its findings

★★☆	Quality rating is “medium”; or
	Quality rating is “high” but high quality external evidence contradicts its findings; or
	Quality rating is “low” but high quality external evidence corroborates its findings
★★★	Quality rating is “high”; or
	Quality rating is “medium” but high quality external evidence corroborates its findings

Rating

★★☆

We estimate that Coffee Initiative II (C.I. II) raised net revenues by $34 per farmer for every $1 that TechnoServe spent on the program. The evidence behind our estimate is of medium quality.

Most of the evidence that undergirds our estimate comes from TechnoServe’s own study of C.I. II. TechnoServe surveyed participating and non-participating farmers in Kenya and Rwanda and found that yields of participating farmers increased substantially more than those of non-participants after receiving the program.[9,10] While we applaud TechnoServe’s efforts to ensure the comparability of participants and non-participants surveyed, we are concerned about several other aspects of the study. The study took place in Kenya and Rwanda, whereas most C.I. II farmers were from Ethiopia. Farmers in Ethiopia were new to the practice of “rejuvenation” — cutting the coffee tree at the base every eight to 10 years, after which it takes another two years to start bearing fruit again.[16] Meanwhile, farmers in Kenya and Rwanda had practiced rejuvenation well before TechnoServe’s intervention. Extrapolating the findings from Kenya and Rwanda therefore does not capture the right pattern of growth in yields for Ethiopian farmers, possibly overstating yield increases in the first two years and understating yield increases in future years. In addition, we suspect farmers who opted into the program were different from those who did not in important ways, such as their level of motivation. This self-selection bias likedly caused the survey results to overstate increases in yields during the two years of observation.

C.I. II has outstanding results compared to other “farmer field school” interventions that have been studied in the research literature. A recent high quality systematic review

found farmer field schools increased net revenues by an average of 19 percent above what participants would otherwise have earned.[17] By contrast, we estimate C.I. II increased net revenues by 48 percent. The large difference is due, in part, to the 10-year horizon over which we extend benefits, based on direct evidence from the program, whereas the large majority of studies included in the systematic review were short term (limited to two years post-training). We commend TechnoServe for its long-term tracking of participants seven years after they entered the program (or five years after exit) — a far longer horizon than has generally been studied in the research literature.

Review

RESULTS OF TECHNOSERVE'S STUDY ON YIELDS

Our estimate is based largely on a "quasi-experimental" study that TechnoServe conducted on C.I. II. In 2012, TechnoServe surveyed close to 300 farmers in Kenya, more than half of whom were new participants in C.I. II who had attended the first few trainings offered.[9] The remaining farmers surveyed served as a comparison group. TechnoServe ran a parallel survey with 300 farmers in Rwanda, also divided into participants (the treatment group) and a comparison group.[10] All 600 farmers were asked to sign a contract promising to record the amount of coffee harvested each day over the course of the study in exchange for a weighing scale.[18] Farmers were given paper calendars on which to record their daily coffee harvests and TechnoServe would visit the farmers every month to collect the previous month's completed calendar and ensure farmers were recording all the coffee they produced. In addition, TechnoServe physically counted the number of productive coffee trees on each farmer's farm every season.

The study enabled TechnoServe to conduct a "difference-in-differences" comparison, comparing the difference in yields (kilograms of coffee produced per tree) of treatment-group farmers from the start of the program (2012) to one year into the program (2013), against the difference in yields for comparison-group farmers over the same timeframe. TechnoServe also gathered data on number of trees on the farm.

The study found higher yields for treatment-group farmers relative to comparison-group farmers.[i] In Kenya, treated farmers harvested 0.88 kilograms more coffee per tree, or a 67 percent increase above the comparison group's yield after the program. In Rwanda, treated farmers harvested 0.79 kilograms more coffee per tree, a 41 percent increase

[i] Statistically significant at a confidence level of 95 percent.

over the yields of their comparison-group counterparts after the program. We use the weighted average of these results, 0.84 kilograms of coffee per tree, in our estimation of C.I. II's impact.

QUALITY OF TECHNOSERVE'S STUDY ON YIELDS

The difference-in-differences study was of medium quality.

TechnoServe's study was a good next-best alternative to a randomized controlled trial. In a randomized trial, farmers from the same population would have been randomly assigned to receive the intervention or not, ensuring that treatment-group farmers were (nearly) identical to control-group farmers (but for statistical noise). Short of running a randomized trial, TechnoServe took other measures to ensure farmers were comparable. Comparison-group farmers were selected from areas where Farm College was not being offered, but that shared similar agronomic conditions to areas where it was on offer, including altitude, soil characteristics and rainfall. TechnoServe made sure that comparison farmers were not directly neighboring treatment farmers to avoid "contamination" of the comparison group, whereby some comparison farmers pick up techniques from newly trained treatment farmers (or even attend training sessions directly). Farmers were then randomly selected to be surveyed from lists of all treatment and comparison farmers. At the first survey in 2012, the average yields of the treatment group and comparison group were similar.

We downgrade the Kenya and Rwanda difference-in-differences study from high to medium quality for two key reasons. First, we suspect it suffers from selection bias. TechnoServe offered Farm College training to all farmers who were members of certain coffee cooperatives. Member-farmers then self-selected into Farm College training.[3] Only about half of all member-farmers took up the training,[8] and it is possible that they were inherently different from their peers who did not. They may have been more motivated, innovative and risk-seeking, perhaps adopting other yield-boosting practices beyond those taught in Farm College. Comparing farmers who self-selected into training against farmers randomly selected from the comparison population might overstate the impact of the program.

Second, the farmers surveyed in Kenya and Rwanda were not representative of the total population of farmers who participated in C.I. II, three-quarters of whom were based in Ethiopia. (This is not a design flaw of the Kenya and Rwanda studies, which represent well the farmers who participated in the two countries. But for our purpose — estimating the impact of the C.I. II program as a whole — the studies are an imperfect source of evidence because of important differences between farmers in Ethiopia and those in Kenya and

Rwanda.) In Ethiopia, coffee trees tended to be far older, in far more need of "stumping" — cutting the tree down close to its base, leaving it out of production for about two years before it starts bear more and better fruit in future years. TechnoServe was largely responsible for introducing stumping (also known as "rejuvenation") to farmers in Ethiopia, training them to stump about one-quarter of their farms each year. In contrast, farmers in Kenya and Rwanda had practiced stumping well before TechnoServe's intervention. If the same surveys in Kenya and Rwanda had been conducted in Ethiopia, TechnoServe would most likely have found smaller increases, if any, in yields: one-quarter of Ethiopian farmers' trees were stumped and out of production at the end of the intervention, but not so at the beginning of the intervention.

On the other hand, the extrapolation of the two-year results from Kenya and Rwanda to the Ethiopian context could also understate the true impact of the program. After two years, the stumped trees in Ethiopia will begin to bear fruit again. Relative to pre-intervention levels, yields are expected to rise substantially. The literature on the rejuvenation of coffee trees predicts large increases in yields until about eight years after stumping, at which point trees should be stumped again — though empirical research quantifying long-term increases is lacking.[16,19,20] We therefore think the two-year results from Kenya and Rwanda do not reflect the full gains that farmers in Ethiopia will see several years later. (For an alternate analysis that assumes an eight-year model of the coffee rejuvenation cycle, please see the annex of this impact audit report.)

DURATION OF BENEFITS

Next, we ask how long benefits persist past the two-year mark of TechnoServe's yield surveys. We lack direct data on how long increases persist. Instead, we look at catch-up in best practice adoption by the comparison group to the treatment group. Said differently, how long it takes for members of the comparison group to adopt the best practices that are boosting the treatment group's yields. To do so, we look at the level of best practice adoption among treatment-group farmers after the program and compare it to their level of adoption seven years after the program, available from an independent assessment by Triple Line. Then, we use difference-in-differences studies on best practice adoption from TechnoServe to estimate how quickly comparison group farmers are likely to pick up new practices. When those two trend-lines meet is when we assume the yields of treatment and comparison group famers converge

We assume, based on Triple Line's independent assessment of the first phase of the Coffee Initiative and TechnoServe's difference-in-differences studies in Ethiopia, Kenya, Rwanda and Tanzania,[11] that coffee yields return to pre-intervention levels 10 years after

the start of the intervention. Said differently, the treatment group sustains a positive difference in yields above the comparison group for 10 years. After then, yields of the treatment group fall back down to meet those of the comparison group.

We consider the independent assessment by Triple Line to be of medium quality. Triple Line acknowledges that the purpose of its study was not to estimate the impact of the program. To do so, it would need some measure of counterfactual effects — what would have happened to farmers had they not participated in the program. Instead, Triple Line set out simply to determine whether participating farmers continued to adopt the best practices they were taught by TechnoServe seven years after the intervention.

To that end, Triple Line used "multi-stage, stratified random sampling" to select 620 farmers in Rwanda to be surveyed. First, it selected three out of 10 cooperatives with which TechnoServe had worked. It is unknown whether the three were selected randomly. Next, Triple Line randomly selected farmers from the three cooperatives. In randomly selecting farmers, Triple Line made sure they were representative of the total list of farmers in terms of their gender and attendance rate at training sessions. The final sample of 620 farmers was large enough to ensure results were statistically significant at the 95 percent confidence level.

We believe Triple Line and TechnoServe chose an appropriate measure of best practice adoption: visual verification in the field by trained surveyors. In contrast to self-reported adoption, visual checks are more objective and likely more accurate. The survey had a fairly high response rate of over 80 percent. However, non-respondents and respondents were different: non-respondents tended to have worse attendance rates. We therefore think the survey results might overstate the average rates of best practice adoption among all participating farmers — some of whom we expect had high attendance, and others, low.

The primary reason we consider the study to be medium quality is its lack of a comparison group. Without a comparison group, the study fails to account for increases in best practice adoption that would have happened in the absence of the program. Adoption of practices, independent of TechnoServe, could have happened due to changes in national agricultural regulations and access to similar interventions delivered by the government or other organizations — both cases that TechnoServe observed in Tanzania.[21]

Therefore, to estimate how quickly a comparison-group farmers would catch up to treatment-group farmers, we turn to TechnoServe's difference-in-differences studies of best practice adoption among treatment and comparison farmers in Ethiopia,[14,22] Kenya,[23]

Rwanda[24] and Tanzania.[21] We draw from only the comparison groups in those studies to estimate the rate at which they adopt best practices without exposure to the intervention.

TechnoServe's 2015 difference-in-differences study of best practices in Ethiopia[14] was well-designed — though we are unsure if the same is true for the studies in Kenya, Rwanda and Tanzania.[i] In Ethiopia, TechnoServe attempted to overcome the non-random assignment of farmers into treatment and comparison groups using "propensity score matching." Recall TechnoServe did not randomly select the regions where it would offer C.I. II. And in the regions where it did offer the program, participants self-selected into the agronomy training. Consequently, it is likely that participants and non-participants differ in meaningful ways, causing different rates of best practice adoption. TechnoServe therefore used propensity score matching to ensure treatment-group and comparison-group farmers matched along key characteristics expected to predict participation: farmer's age; farm size; roof material;[ii] and the number of best practices adopted *before* the intervention. Sample sizes were adequate to detect the effects of the intervention at the 95 percent confidence level.

Using this evidence, we construct trend-lines for the level of best practice adoption after the program among the treatment and comparison groups. Those lines converge at 10 years, which we take as the duration of the increased yield for treatment-group farmers.

PRICE PREMIUM

Farmers benefit from C.I. II in two ways: higher yields and higher selling prices of "wet-milled" coffee. Wet-milled coffee has been pulped, fermented, washed and dried in a wet mill. Wet-milled coffee is a specialty item because it is often more consistent in quality and has a more balanced flavor profile than regular sun-dried coffee.[2]

TechnoServe field staff recorded the prices of wet-milled and sun-dried coffee in the countries where C.I. II was implemented. Staff consulted local traders and verified that their answers matched what farmers claimed traders were paying.[3] Triple Line worked closely with TechnoServe's monitoring and evaluation staff in Ethiopia to collate the price data for their independent assessment. We, in turn, use the price data as reported in Triple Line's evaluation, making no adjustments. However, without more information on how this data was collected (from whom, how often, and using what data-collection instruments), we are unable to judge its quality. TechnoServe has a long history in coffee

[i] Based on documents received from TechnoServe, this methodology was only used for the 2013 cohort in Ethiopia.

[ii] The building material of one's roof is commonly used as a predictor of a household's roof, for instance by the Progress out of Poverty Index.

agronomy. We expect that large discrepancies in price data would have been caught by the experienced TechnoServe team or by their evaluation partner, Triple Line.

COMPARISON TO THE RESEARCH LITERATURE

Our estimate of the impact of C.I. II is far larger than estimates of similar farmer training programs that have been studied in the research literature. C.I. II boosted the net revenues of participants by about 48 percent above those of non-participants. The average effect size reported by the best empirical studies of farmer field school interventions is 19 percent, according to a systematic review.[17]

The discrepancy is large but expected. Our model counts 10 years of impacts, based on evidence and educated assumptions, whereas the large majority of studies in the literature are limited to two years post-training. Farmer field schools also tend to be more effective at increasing net revenues earned from cash crops such as coffee. The systematic review covered interventions for both subsistence crop farmers and cash crop farmers, so it is expected that C.I. II achieved above the average effect size reported in the review. Further, the systematic review identified elements of intervention design and operations that tend to boost the impact of farmer field schools, such as training delivered in a participatory manner, strong social networks among participants and concrete demonstration of the benefits of best practices being promoted. If TechnoServe's program incorporated all such "enablers" of success, we would expect it to achieve better than average results.

ANNEX

Nonprofit Information

NAME	TechnoServe
CHARITABLE STATUS	501(c)3 nonprofit
WEBSITE	www.technoserve.org
CONTACT EMAIL	info@technoserve.org
ADDRESS	1120 19th Street NW, 8th Floor Washington, DC 20036

Audit Information

RELEASED	
PERMALINK	www.impactm.org/a/technoserve/1
STANDARD	Version 0.3
ACTIVITIES	Literature review, document and data review, senior management interviews, field staff interviews and key informant interviews.
AUDIT TEAM	Tamsin Chen and Ben Mazzotta
REVIEW TEAM	Elijah Goldberg and Michael Weinstein
CONFLICT DISCLOSURES	None

Alternate Approach to Analysis of Impact and Cost

The benefit/cost ratios presented in the impact audit could be overestimates or underestimates. Indeed, the upward and downward biases might even "cancel" each other out. In this supplemental analysis, we identify the key analytical decisions made in the audit that would tend to overstate or understate results and test what happens if we vary them.

CONTROLLING FOR DIFFERENCES BETWEEN PARTICIPANTS AND NON-PARTICIPANTS

Our analysis relied on a study by TechnoServe that compared farmers who volunteered to participate in the program to farmers who did not. If farmers who volunteered were, for instance, more motivated and resourceful than non-volunteers, then the study likely attributed too much impact to the program — impact that volunteers would have achieved on their own, without TechnoServe's intervention. Here, we explore what would happen to our estimates if we instead used a different source of evidence that did control for inherent differences between participants and non-participants: a randomized controlled trial (R.C.T.) conducted by third-party evaluators on the first phase of the Coffee Initiative in Rwanda (not the second phase, which is the subject of our audit).[25] The R.C.T suffered major problems of its own, hence we do not use it as the basis of our primary results.[i] But it did correctly use randomized assignment (and statistical controls), ensuring participants and non-participants were comparable. Farmers signed up to participate in the program, then *of that list of farmers*, the evaluators randomly assigned some farmers to receive the program and some farmers to serve as the control group. Using the R.C.T. results as the basis of our estimates gives a benefit/cost ratio of 25:1, counting only TechnoServe's costs, compared to 34:1 originally. The benefit/cost ratio counting only farmers' costs is 52:1, compared to 71:1 originally.

[i] The baseline and first-year surveys required farmers to self-report their yields from memory. For the baseline survey, farmers had to recall a whole year's worth of coffee harvested. Surveys during the first year of the intervention required farmers to recall harvests from at least four months at a time.[25] For the second and third years of data collection, the trial used the same paper calendar and weighing scale method as the difference-in-differences study described in the Quality of Evidence section. But before they were issued calendars and scales, farmers previously did not keep any records. Program staff say any information reported from memory is likely so inaccurate as to be unusable, especially because coffee fruit is harvested continually rather than in infrequent (and therefore memorable) instances.[3]

ADJUSTING FOR THE "MONITORING EFFECT"

There is evidence that treatment-group farmers exhibited a "monitoring effect," as observed in Laterite Ltd.'s independent assessment of the first phase of the Coffee Initiative.[18] TechnoServe's team of data collectors and supervisors visited farmers every month during the harvesting season to ensure that farmers were recording all the coffee they produced. They also counted the number of fruit-bearing trees on each farm. This frequent interaction and surveillance may have caused the 11 percentage point increase in attendance at training sessions and five percentage point increase in best-practice adoption among *surveyed* treatment-group farmers relative to *non-surveyed* treatment-group farmers.[26] The differences were statistically significant. Important, differences in attendance only started to become apparent when the yield surveys began. Meanwhile, even though comparison-group farmers selected for the yield survey also received the same attention from TechnoServe data collectors, that extra attention would not have manifested in more diligent attendance or best-practice adoption. Those farmers simply did not have access to the intervention at all. They might have altered their behavior somewhat because they were being observed, but likely not in the specific ways encouraged by Farm College training. Adjusting the R.C.T.-based results for the "monitoring effect" further lowers our benefit/cost ratios to 24:1, counting only TechnoServe's costs, and 50:1, counting only farmers' costs.

THE COFFEE REJUVENATION CYCLE

We would obtain yet another set of benefit/cost ratios if we followed the agricultural *theory* of coffee rejuvenation. Agronomy manuals and sector experts agree that after a coffee tree is "stumped" (cut at its base to enable greater yields in the future), the tree bears no fruit for up to two years, after which its yields improve dramatically for six years.[16,19,20] After a total of eight years, the tree must be stumped again. Farmers participating in Coffee Initiative II were taught to stump one-quarter of their coffee farms each year so that the remaining trees would continue bearing fruit and providing revenue.[8]

The problem: we could find no empirical evidence in the literature on the long-term magnitude of yield improvement from stumping. We therefore limited our original model to the two-year results from TechnoServe's own study of yields, potentially missing out the future gains from stumping.

However, TechnoServe program staff have created a conceptual model of yield increases after stumping. The model assumes that the yield of a recently stumped tree falls to negative 10 percent of its original volume in the first year after stumping; negative 20 percent in year two; zero in year three; 25 percent in year four; and increases gradually

until it reaches 100 percent of its original volume in year eight. We combine this model with survey results from Triple Line and TechnoServe that show the percentage of farmers that adopt rejuvenation at the baseline, two-year and seven-year marks. The result is a large increase over our original estimates: a benefit/cost ratio of 51:1 from the perspective of TechnoServe, or 104:1 from the perspective of farmers.

Benchmarking Impact and Cost Estimates

Taking the perspective of TechnoServe, ImpactMatters estimates that Coffee Initiative II increased the net revenues of the average participating business by $34 for every $1 that TechnoServe spent on the program (a benefit/cost ratio of 34:1). Each business earned an additional $3,300 in profits over 10 years.

Summarizing the results of their systematic review,[17] Waddington and White discussed the cost-effectiveness of "farmer field school" interventions like Coffee Initiative II, drawing from four cost-effectiveness studies, as well as cost data from a set of 337 interventions.[27] They found farmer field school interventions cost implementing organizations an average of $60 per farmer, though most cost between $20 to $40 per farmer. These figures do not take into account the costs to farmers of attending trainings and implementing labor-intensive practices. The equivalent cost in TechnoServe's case is $100.

Though TechnoServe's cost appears high by comparison, it has a far more favorable benefit/cost ratio than any of the interventions reviewed by Waddington and White. The highest benefit/cost ratio they found was 6.8 — for program that trained farmers in Bangladesh on integrated pest management (I.P.M.). The lowest: 0.42 for a program in China that trained cotton farmers in I.P.M. Meanwhile, we estimate TechnoServe's benefit/cost ratio at 34:1 (counting only costs spent by TechnoServe to deliver the program).

There are several likely reasons for the large difference. First, besides training farmers, TechnoServe also introduced wet-mill technology to farmer cooperatives, enabling farmers to sell their coffee at higher prices. Second, TechnoServe observed continued adoption of best practices five years after the program concluded, allowing us to model far more years of impact than have generally been studied in the research literature. Third, Waddington and White note the effect of farmer field schools on net revenues is particularly strong for cash crops like coffee. Fourth, self-selection bias may have biased upward the benefit/cost ratio, as discussed in the Quality of Evidence section. Finally, there may simply have been features of TechnoServe's intervention — the right programmatic design and strong personnel, for instance — that made it particularly successful.

Glossary

Bias

Bias is a non-random error in a statistical estimate. Whenever estimates are based on a sample from a larger population, there will be random error in that estimate: no two samples will produce exactly the same estimates. An estimate is biased when those errors lead it to be consistently above or below the true value that is being estimated.

Comparison Group; Control Group

A comparison group, in contrast to the treatment group, is a group that did not receive the intervention. Comparison groups enable nonprofits and researchers to compare what happened to participants of their program to what might have happened if they were not in the program. ImpactMatters refers to comparison groups as "control groups" if they were constructed using probabilistic sampling, meaning if control-group members were chosen at random from the same population as the treatment group.

Counterfactual; Counterfactual Evidence

The counterfactual is what would have happened in the absence of a program or other event. Understanding the counterfactual is essential to understanding the impact of a program. Participant outcomes may change over time for many different reasons not related to the program. By comparing the difference between participant outcomes and counterfactual outcomes, the impact of a program can be estimated.

The counterfactual cannot be directly measured, as researchers cannot observe the same participant both participating and not participating in the program. However, it can be approximated by randomizing participants into an intervention group and a control group, and then comparing outcomes across the two different groups.

Difference-in-differences

A statistical technique that compares the change over time in the outcome variable of the treatment group, to the change over time in the outcome of the comparison group. It may be used for multiple time periods and multiple groups. Common variations of the term include "difference in difference" and "difference-in-difference."

Discount Rate

People tend to value benefits in the future less than benefits in the present, for three primary reasons. First, benefits today can be reinvested and generate some return. Second, the future is uncertain, and we are often uncertain if future benefits will actually

materialize. Third, most people are impatient, and prefer immediate gratification over future gratification. A discount rate captures this by discounting or reducing future benefits compared to current benefits.

Effect Size

How large the measured impact was on outcomes in the group receiving the program compared to a similar group that did not receive the intervention.

GRADE

Grading of Recommendations Assessment, Development and Evaluation (GRADE) is an approach to rating the quality (or certainty) of evidence and strength of recommendations. ImpactMatters' assessments of quality of evidence are inspired by the GRADE approach.

Impact

Impact is a change in beneficiary outcomes attributable to a nonprofit’s intervention, net of counterfactual effects.

Independent Evaluator

An independent evaluator can include a research organization or academics engaged to analyze the impact of a program. Independent evaluators are not directly employed by the program, although they may be paid through program resources.

Intervention

An intervention is what researchers study and nonprofits implement. An intervention includes anything from a medical procedure to a conditional cash grant. ImpactMatters studies the intervention that a nonprofit implements, mapping that intervention to the evidence base on that particular intervention. Also referred to as the nonprofit’s program.

Purchasing Power Parity

The purchasing power of a currency is the quantity of the currency needed to purchase a common basket of consumer goods and services. P.P.P. equalizes the purchasing power of two given currencies by accounting for differences in the cost of living and inflation in the two countries.

Quality of Evidence

Quality of evidence captures ImpactMatters' confidence in our impact and cost estimates. For programs with high-quality evidence, the impact and cost estimates are more likely to

accurately reflect the effectiveness of the program. Quality of evidence reflects only the data used to construct the impact and cost estimate. It is rated using an adaptation of the GRADE methodology, a systematic approach to judging evidence.

High-quality evidence under the GRADE rubric is the best scientific evidence that the program has its intended impact. Randomized designs are presumed to be in this category unless our analysts are concerned about flaws in the methodology or weak results.

Medium-quality evidence under the GRADE rubric has some flaws that might render estimates of impact inaccurate. Quasi-experimental designs are presumed to be in this category unless flaws are mitigated and results are convincing. Those designs can also be rated down to low quality if our analysts are concerned about the methodology or results.

Low-quality evidence under the GRADE rubric limits our confidence in the estimate of impact. Observational studies are presumed to be of low quality unless flaws are mitigated and the research shows very convincing results, such as with a large effect size and a clear dose-response curve.

Very-low-quality evidence under the GRADE rubric gives us very little confidence in the estimate of impact. Flawed observational studies, and even quasi-experimental or experimental studies with multiple, serious flaws, might fall into this category.

Quasi-experimental Design

A study with a quasi-experimental design tests a causal hypothesis, but lacks random assignment of test subjects to treatment and control groups, perhaps due to logistical or ethical constraints.

Randomized Controlled Trial (R.C.T.)

A randomized control trial is an evaluation design by which individuals (or groups) are randomly allocated into treatment and control groups, where the treatment group receives the program. The outcomes of the two groups are then compared in order to estimate effect size.

Sample; Sample Size

The sample is the portion drawn from a population for testing or analysis that is intended to enable statistical estimates of the behavior or attributes of the whole population. The sample size is the number of units that comprise the sample; a large enough sample size allows inferences about the whole population to be made.

Social Costs or Societal Costs

Social costs include all costs incurred by society as a result of the nonprofit's program. Different from accounting costs, which include just the costs that appear on the nonprofit's accounting statements, social costs may include, for instance, the opportunity costs of participants' time spent in the program and the costs to other organizations and governments of helping to delivering the program.

Statistical Power

Statistical power is the probability that a test will correctly reject the null hypothesis (the hypothesis that there is no statistically significant difference between the samples being compared). An underpowered test will likely yield large p-values and confidence intervals, and will lack the evidence to reject the null hypothesis.

Statistical Significance

A statistically significant result (often a difference of means of the main outcome of interest) is a result that is unlikely to arise as a result of chance. This doesn’t mean the finding cannot be due to chance – just that it is very unlikely.

Systematic Review

A type of literature review that collects and analyzes multiple research studies in order to answer a research question. After a research question is defined and appropriate research studies identified, data from the studies are extracted, assessed for their quality, analyzed, sometimes statistically combined in meta-analyses, and reported in such a way as to address the research question.

Theory of Change

A theory of change connects the problem to the intervention the nonprofit runs to expected process and outcome metrics. The objective of a theory of change is to provide a testable hypothesis for why the intervention is solving some problem that will lead to positive changes for the targeted beneficiaries.

Treatment Group

In an experiment, the treatment group is comprised of experimental subjects that receive the treatment being evaluated.

References

1. BMGF.Grant.Agreement.PhaseII.2011.11.16.
2. Brewing Prosperity in East Africa [Internet]. [cited 2017 Jul 4]. Available from: http://www.technoserve.org/files/downloads/Coffee-Initiative-Final-Report.pdf
3. S0247 Interview.
4. BMGF005N Coffee Initiative Phase 2 Progress Report 4 FINAL.
5. Coffee Initiative Phase 2 Progress Report 2_submitted.
6. de Mel S, McKenzie DJ, Woodruff C. Measuring microenterprise profits: Must we ask how the sausage is made? J Dev Econ. 2009;88(1):19–31.
7. Evaluation of the TechnoServe East Africa Coffee Initiative. 2017 [cited 2017 Jul 5]; Available from: http://www.technoserve.org/files/downloads/triple-line-evaluation-of-the-coffee-initiative.pdf
8. S0249 Emails with TechnoServe Staff.
9. Appendix 9.1. Coffee Yield Midline Scorecard (2012 Cohort - Kenya_ Provisional).
10. Appendix 9.2. Coffee Yield Midline Scorecard (2012 Cohort - Rwanda).
11. IPE Triple Line. Evaluation of the TechnoServe East Africa Coffee Initiative [Internet]. 2017. Available from: http://www.technoserve.org/files/downloads/triple-line-evaluation-of-the-coffee-initiative.pdf
12. International Coffee Organization. The Current State of the Global Coffee Trade [Internet]. 2016 [cited 2017 Dec 1]. Available from: http://www.ico.org/monthly_coffee_trade_stats.asp
13. Appendix 14. Maize Learnings Report (Ethiopia).
14. Appendix 1. Coffee Best Practice Endline Analysis (2013 Cohort - Illubabor_ Ethiopia).
15. Appendix 15.1_120215 EAFCA Sustainability Forum Presentation.
16. Dalberg Advisors. Renovation & Rehabilitation for Resilient Coffee Farms: A Guidebook for Roasters, Traders and Supply Chain Partners [Internet]. 2017. Available from: https://www.conservation.org/publications/Documents/RR Guidebook_vfinal.pdf
17. Waddington H, Snilstveit B, Garcia J, Vojtkova M, Anderson J. Farmer field schools for improving farming practices and farmer outcomes in low- and middle-income countries : a systematic review. Campbell Syst Rev. 2014;(6):336.

18. Sabates R, Gathani S, Stoelinga D, Gomez MP. Independent Assessment of TechnoServe's Coffee Agronomy Training Program [Internet]. 2013 [cited 2017 Jul 4]. Available from: http://www.technoserve.org/files/downloads/laterite-independent-agronomy-assessment.pdf

19. Netsere A, Kufa T, Tesfaye S. Review of Arabica coffee management research in Ethiopia. J Biol Agric Healthc. 2015;5(13):235–58.

20. Netsere A, Shimber T, Kufa T, Taye E, Gebreselassie W. Yield Response of Forest Arabica Coffee to Ridges and Rejuvenation Methods. In: 21st ASIC International Conference on Coffee Science. Montpellier: Association for Science and Information on Coffee; 2006.

21. Appendix 8.3. Coffee Best Practice Endline Analysis (2012 Cohort - Tanzania).

22. Appendix 1. Coffee Best Practice Endline Analysis (2012 Cohort - Jimma_ Ethiopia).

23. Appendix 8.1. Coffee Best Practice Endline Analysis (2012 Cohort - Kenya).

24. Appendix 8.2. Coffee Best Practice Endline Analysis (2012 Cohort - Rwanda).

25. JPAL_2012_Coffee_Rwanda_Internal.

26. Gathani S, Gomez MP, Sabates R, Stoelinga D. The Effect of Monitoring: How Data Collection Type and Frequency Boosts Participation and the Adoption of Best Practices in a Coffee Agronomy Training Program in Rwanda. Eval Rev. 2015;39(6):555–86.

27. Waddington H, White H. Farmer field schools: From agricultural extension to adult education. 3ie Systematic Review Summary Reports. 2014.