Scaling Skills[*]

Maria Sofia Casabianca,[†] Megan Lang,[‡] and Julia Seither[§]

December 30, 2024

Abstract

This paper evaluates whether skills-based programs can scale cost effectively by training local non-government organizations as implementers. Using two field experiments in Uganda, we compare the impacts of an entrepreneurship and business skills program when delivered by original program coaches versus local NGOs trained in the curriculum. Results show no significant differences in business outcomes or skills between the two types of implementers. Training local NGOs reduces costs per participant by 45% due to lower local implementation expenses. These findings highlight decentralized scaling as a viable method to preserve program impacts while improving cost-effectiveness, addressing scalability challenges in development interventions.

KEYWORDS: Scaling, Skills, Firms, Uganda.

JEL CODES: D13, J24, L39.

1 Introduction

Programs implemented at scale often have smaller impacts than those implemented by researchers during trial periods (Vivalt 2020, Araujo, Rubio-Codina, and Schady 2021). This "scale-up effect" stems from multiple causes: different populations and contexts, loss of implementation fidelity, and spillovers, among others (Al-Ubaydli, Lee, et al. 2021, Al-Ubaydli, List, and Suskind 2019, Banerjee, Banerji, et al. 2017). Given these challenges, we consider an alternative path to scale: teaching *local* non-government organizations (NGOs) how to implement an effective program.

We experimentally test the implementation and cost-effectiveness of a skills-based program when implemented by the original organization who developed it compared to when implemented by other local NGOs. The program we consider teaches business and entrepreneurship skills. It was originally designed to serve ultra-poor, low-literacy women in Uganda. Developing skills is an integral part of numerous anti-poverty programs (e.g., BRAC, Concern International), but bringing skills-based programs to scale poses unique challenges. Ability and motivation among both instructors and participants may vary more as programs scale. Furthermore, if skills-based programs rely on individual mentoring or small class sizes, costs will rise somewhat linearly as programs scale. Conversely, skills-based programs may be well-situated to scale in a decentralized manner precisely because they primarily rely on staff time, which may be easier to reallocate to new programs than other organizational resources. Local staff may also have existing relationships with participants or a better understanding of local communities that make them more effective. As such, it is unclear how well skills-based programs can be scaled through decentralized implementation.

The original program we study is called Street Business School (SBS).[1] Over the course of six months, program coaches teach hard and soft skills for entrepreneurship through eight

---

[1] The original NGO that developed and implemented the program is also named Street Business School (SBS). In the following we refer to SBS locations as those locations where the original NGO implemented the program and to GCP locations (Global Cooperation Partners) as those where local NGOs that have been trained on the program implement it.

classroom modules and individual coaching sessions. The program is purely skills-based, as it offers no direct capital or financing to participants. In addition to directly implementing their program, SBS runs a global expansion program in which they train staff from other NGOs to teach their curriculum. Although coaches trained through the expansion program have less experience teaching the SBS curriculum than SBS coaches, they often have well-developed relationships with the communities where they work and already run programs that potentially complement the skills taught in SBS. Thus, coaches trained through the expansion program are less experienced with the SBS curriculum, but have skills and local knowledge that could enhance the curriculum's effectiveness.

We assess the potential for scale by comparing the results from two randomized control trials (RCTs). The first RCT measures the effectiveness of SBS when it is implemented by SBS coaches. The second evaluates the program when implemented by coaches of three organizations trained through the expansion program. Although both RCTs study program implementation in Uganda, the population each organization works with differ. We discuss these differences in detail in section 2.

Our three primary results show that coaches trained through the expansion program are as effective as SBS coaches. First, we find no significant differences in effects on business performance between the two RCTs. Second, effects on hard skills (business tracking, price management, and goal setting) show no significant differences between the two RCTs. Third, we find similar patterns for soft skills, although qualitatively it appears that the specific soft skills impacted may differ between groups coached by SBS versus expansion partners.

Having established that there are no declines in program impacts when the SBS curriculum is implemented by expansion partners, we turn to the question of cost-effectiveness. Based on estimates provided by SBS and selected expansion partners, we estimate that the total cost per participant for expansion partners is just 55% the cost for SBS, even accounting for higher training costs for expansion partners relative to SBS coaches and assuming that expansion partners will implement the curriculum with fewer participants over the course of

3

their careers. Lower implementing costs for expansion partners relative to SBS drive the cost reductions. Taken together, our results suggest that the expansion model provides a path to scale that preserves program impacts and improves program cost-effectiveness.

Our work speaks to two related strands of literature on scaling: studies evaluating the same intervention in different contexts and evaluations of interventions at scale. As Duflo (2004) points out, evaluations of the same intervention in different contexts speak to external validity, a key requirement for scalability. Our work joins that of Banerjee, Duflo, Goldberg, et al. (2015), which evaluates multifaceted poverty reduction programs in different contexts, unified studies of micro-credit (Banerjee, Karlan, and Zinman (2015), Banerjee, Duflo, Glennerster, et al. (2015), Tarozzi, Desai, and Johnson (2015), Attanasio et al. (2015), Crépon et al. (2015), Angelucci, Karlan, and Zinman (2015), Augsburg et al. (2015)), and education programs (Bando, Naslund-Hadley, and Gertler (2019)). Although our work focuses on a single country, we study the effects of a specific, skills-based intervention implemented by different organizations and with different populations. Our work shows that skills-based interventions can scale through a training of trainers model even when the trainers and the final populations being served differ substantially.

Evaluations of interventions at scale typically focus on scaling from researcher or NGO-led implementation to government implementation, often with smaller effects (Banerjee, Banerji, et al. (2017) is an important exception). For instance, Araujo, Rubio-Codina, and Schady (2021) discuss the declining effects found when scaling parenting home visits from Jamaica to larger programs in Colombia and Peru. In a large meta-analysis, Vivalt (2020) shows that government-led programs have consistently smaller effects than programs implemented by researchers or NGOs. Importantly, she finds that the relationship holds even when accounting for sample size, highlighting the importance of building alternative paths to scale. We study an alternative path to scale that continues to leverage the expertise of local NGOs rather than relying on governments. Our results provide a proof of concept for the decentralized scaling of a program through local organizations, establishing that programs can scale through the

existing ecosystem of NGOs without eroding program effectiveness.

We proceed by describing the context of both the original RCT with SBS coaches and the RCT of expansion partners, along with a detailed discussion of the populations participating in each study. We then describe the curriculum before presenting the experimental design and empirical results on program impacts in the two studies, then discuss costs. We conclude by discussing the implications of our work and directions for future research on alternative models of scaling.

2 Background and Context

Our results on the Street Business School (SBS) curriculum as implemented by SBS coaches come from an RCT conducted in five communities in central Uganda between 2018 and 2021. SBS selected the five locations in the same way they would typically select a new community: by speaking with community leaders, assessing the economic needs in the community, and determining whether a sufficient number of women would be interested in participating. Of the five communities in the initial RCT, four are rural and one is peri-urban. This RCT included two treatment groups. Both followed the same classroom curriculum, but one had optional mentoring at the training site while the other had "mandatory" mentoring at a woman's home or business (although there is still imperfect compliance to the mandatory mentoring). The version with mandatory mentoring is the one taught to expansion partners, so we show results both including and excluding participants from the optional mentoring group.[2]

From November 2019 to April 2023, we worked with three non-government organizations (NGOs) in Uganda who had previously taken part in Street Business School's (SBS) Global Cooperation Partners (GCP) program. Unlike SBS coaches, expansion partners have existing programs running in communities, so they run Street Business School within communities

---

[2]The pre-analysis plan cited above describes the full design of this research project. The full set of results (that go beyond the scope of this paper) can be found in Lang and Seither (2022) and Lang and Seither (2024).

where they are already working rather than identifying new communities. Importantly, we instructed the NGOs who participated in our study to run the program precisely as they usually would. The goal was not to replicate the SBS curriculum as closely as possible but instead to see whether the ways expansion partners implement the SBS curriculum in practice is effective.

In total, we present results from the five cohorts in the first RCT with SBS coaches and three cohorts enrolled through three separate NGOs in the second RCT. Although all three expansion NGOs work in Uganda, the services that they provide apart from the SBS curriculum differ widely. One is largely focused on raising awareness about disabilities, providing therapeutic services for disabled persons, and teaching technical skills to caregivers. One pursues a variety of development projects, including running a health center and school, promoting agribusiness, and running a communal mill. The last is focused on empowering women and girls through vocational programs, education, and programs on the environment and sanitation. Unlike SBS, expansion partners do not always limit participation to women.[3]

Since expansion partners already run multiple programs, a natural question is whether adding the SBS curriculum to their program offerings crowds out other activities. Although we lack direct evidence on changes in expansion partner activities, it is worth noting that expansion partners seek out training from SBS to learn the curriculum. The typical workshop for expansion partners lasts for eight days and costs up to US$4,900 for two participants per organization.[4] Given the costs involved, it appears that expansion partners view the SBS curriculum either as an important complement to their existing programs or as an important new program to implement as their needs change over time.

When SBS coaches implement the program, they recruit cohorts of participants by distributing flyers door to door and hiring local advertisers to make announcements on motorbikes in the days leading up to an orientation event. Women attend the orientation event to

---

[3]From the 1,334 participants we had at baseline only 50 were men, accounting for just 3.7% of our total sample. Given this small proportion of men, we do not evaluate differential treatment effects by gender.

[4]Organizations are required to send two participants. Organizations can receive scholarships funded through donations to SBS, but all organizations must pay at least US$1,500 for two attendees.

learn more about the program. If they decide to participate, they sign up at the orientation event. Recruitment among expansion partners can differ because they have have pre-existing populations they work with in their communities.

Given the number of activities already undertaken by each NGO and differences in the selection process, we do not expect that the communities in our study of expansion partners will be directly comparable to the communities where we evaluated the impact of SBS when delivered by SBS coaches. Examining baseline differences between the two RCTs reveals a number of differences. Although roughly half of participants in all locations have a business at baseline, participants in expansion partner locations tend to own fewer businesses than women in the SBS locations. For those participants with businesses, baseline sales and profits tend to be lower. Demographically, participants in expansion partner locations tend to be younger, to have larger families, and are more likely to be employed at baseline.[5]

3   Experimental Design

3.1   Treatment

The Street Business School curriculum teaches business and entrepreneurship skills with no requirements for literacy or numeracy. The program runs for six months and includes eight classroom sessions of around 2 hours each and three one-on-one mentoring sessions. The first three months focus on starting and running a business and the last three months focus on skills for firm growth. The first module is called "getting out of your comfort zone" and aims to show participants that they have untapped potential. The second is "identifying business opportunities", which helps participants consider business ideas that may be successful in their communities. The third is called "finding capital and starting small". As the program does not provide capital, this module is designed to help participants understand how to raise funds to start a business.

---

[5]Descriptive statistics on the different samples can be found in Appendix Table A1.

During the second month, the program involves two modules on management practices. The first teaches bookkeeping and record keeping using techniques similar to the "rules of thumb" in Drexler, Fischer, and Schoar (2014). The second is designed to help participants do "market research" to understand the local market before investing their time and resources to start a business. The third month only has one module: business planning. In this module, coaches show participants the steps to planning a business and emphasize the benefits of developing a plan before trying to start a new business.

Month four of the program has two modules. The first is "growing your customer base", which covers topics like actively pursuing customers, customer service, and offering promotions. The second module is "money management", which teaches the value of saving and budgeting and provides tools to help participants start separating and prioritizing personal versus business expenses. Month five is entirely given to implementation. Ideally, participants start or continue working on their business this month using the skills they have learned. The program ends with a formal, public graduation ceremony to celebrate the achievements of the women who participated. During the six-month teaching period, coaches do their best to schedule three one-on-one, on-site visits at the participant's homes or businesses. These visits are meant to give the participant individualized coaching and respond to questions specific to their business.[6,7]

SBS explicitly teaches that their curriculum should not include access to capital or cash. They teach women to finance their business through savings, formal and informal loans, or by leveraging smaller, less capital-intensive businesses into larger, more capital-intensive ones. The program teaches that even small amounts of money may be enough to start growing an enterprise. SBS considers the lack of capital provision to be critical because the women they work with are often exposed to negative income shocks that can cause their businesses to

---

[6]SBS and expansion partners implement on-site mentoring visits. However, we only have compliance data for the RCT with SBS coaches.

[7]The RCT with SBS coaches also had an opt-in mentoring group. Results between the two groups are not statistically significantly different for most outcomes at midline when the program ends. See Lang and Seither (2022) for results on this group.

close. By teaching women to raise capital rather than providing it, the curriculum attempts to ensure that women can restart their enterprises after such shocks. Although we cannot rule out that expansion partners may have ways of providing access to finance or capital to their participants, this aspect of the SBS curriculum is central to the decentralized scaling model. Partner organizations only need to fund or reallocate staff time to implement the curriculum as intended.

3.2 Sampling Frame

Our sampling strategy maintains the self-selection that typically occurs at the start of SBS. For the RCT with SBS coaches, women self-select into the program after attending an orientation day where they learn about the program. Selection can differ for expansion partners if they are only recruiting among populations they already serve or otherwise targeting their recruitment within the community. As such, our estimated impacts are intentionally inclusive of differences in selection into the program, differences in the populations served, and differences in coaches' abilities to implement the program.

We enrolled a total of 1,334 participants in 8 locations in central and western Uganda over the course of four years. SBS selected the expansion partners we worked with based on their capacity and willingness to meet the sample size needed for the experiment. SBS also selected the five communities to work in during the RCT of SBS coaches.

3.3 Assignment to Treatment

We implemented a double-blind, individual-level randomization at the end of the baseline survey. We gave the enumerators a bag with colored candies, and they asked participants to select one. After the participant had drawn the candy, they received a matching-colored paper with information about the time, date, and venue of the first training session. At no point before the first day of training did we reveal to either the participants or enumerators which color corresponded to each treatment arm. Participants found out about their treatment

9

status during the first session at the venue.[8] The control group did not receive any training during the RCT, but participants in this group met during the first day of the program to ensure compliance with randomization and to provide an opportunity for them to ask the research staff questions regarding when they would be eligible to participate in the program.

We assess baseline balance across various factors, including gender, age, marital status, education level, parental education, employment status, household size, number of minors living in the household, and business ownership in Table A1. We examine selective attrition using the same variables in Table A2. The imbalance we observe is in line with what we would expect by chance given the number of covariates we test. Most of the imbalances appear on educational attainment of the participant and their parents. We control for these in our main specification along with gender, household size, and the number of children in the household. Attrition is correlated with some baseline covariates (see table Table A2). Participants who were married or cohabitating at baseline are more likely to drop from the sample, and those participants who report having a business at baseline are less likely to drop out. There is a marginally significant probability that people in the RCT of expansion partners are less likely to leave the program.

4   Empirical Strategy & Data

We estimate intent-to-treat (ITT) effects using data collected just after the program ended (approximately six months after the baseline survey). Our primary results use an ANCOVA specification of the form

$$O_{it} = \alpha + \beta_1 SBS_{it} + \beta_2 GCP_{it} + \delta_1 X_i + \delta_2 X_i \times GCPLocation_{it} + \gamma_1 O_{i0} + \nu_j + \epsilon_{it}, \quad (1)$$

where β_1 gives the ITT of participating in the program when delivered by SBS and β_2

---

[8]Note that the number of participants in treatment and control is not precisely even as a result of random chance.

gives the ITT of the program when delivered by an expansion partner (GCP). We control for gender, age, household size, number of minors living in the same household, education, marital status, and location fixed effects.[9] We adjust standard errors for multiple hypothesis testing by family of outcomes for each table following Romano and Wolf (2005).

We additionally report results from a larger dataset that includes the second SBS treatment group with voluntary mentoring at the training center.[10] This secondary specification includes a binary indicator for treatment status and a binary variable indicating whether participants were in the mandatory or voluntary mentoring group. Since the program implemented in expansion partner locations follows the original SBS implementation strategy (on-site mentoring at respondent's homes or businesses), the secondary specification is a less precise comparison.

We test whether the training translates into business outcomes in a similar way when delivered by an expansion partner compared to SBS. At the extensive margin of business outcomes we estimate effects on owning a business and earning positive sales revenues and profits. We differentiate between income generated from the main business, other businesses, and the combined profit of all businesses.[11] We combine these measures in a summary index of business performance on the extensive margin. To understand whether businesses of treated participants are more successful than those of participants in the control group, we estimate intensive margin effects. We estimate treatment effects for the number of businesses a respondent owns, sales revenues and profits for the main business, profits from all other businesses, the combined profits from all businesses, and a summary index.

Lastly, we estimate effects on hard and soft skills. Hard skills relate to classical business practices. We collect data on whether respondents track different business metrics (sales,

---

[9]Only 12.6% of participants in expansion partner locations are men, or 3.7% of our total sample.

[10]Preliminary results on treatment differences by mentoring modality can be found in Lang and Seither (2022). The results suggest that implementing SBS with office hours is equally effective right after the program ends - the time frame we consider in this study as well. However, the program with voluntary mentoring seems more effective than with on-site mentoring in the long-run. A detailed analysis on the impact of SBS with office hours can be found in Lang and Seither (2024).

[11]We define the main business as the business earning the highest profits at the time of the survey.

expenditures, and profits), on a price management index, and on an index built upon multiple questions about goal setting for the business. We combine these indicators into a single summary index to assess overall effects on hard skills. Soft skills include self-efficacy, grit, and three measures of locus of control: internal, powerful others, and chance. Again we combine these indicators into a single summary index.

5   Results

The program is equally effective at improving rates of profitable business ownership when implemented by GCP coaches as when implemented by SBS coaches. Columns (1) and (2) of Table 1 show that, compared to the control group, participants are around 14pp more likely to own a business independent of the implementer. We find positive impacts on the likelihood of earning positive sales and profits compared to participants in the control group, but there is no statistically significant difference between implementers: the program increases a participant's likelihood of earning positive profits from their main business by approximately 13pp in both cases. Although the program when implemented by SBS coaches also helps participants to generate profits from other businesses, we find no impact on other business profits for GCP participants, although the difference between the implementers is not statistically significant. In line with these results, we find that the program increases the probability of earning positive profits across all businesses, and we find positive impacts of the program when combining these measures into an index. Although there are no significant differences between treatment groups in the summary index, effects are slightly larger for SBS-implemented groups due to qualitatively larger effects on earning positive profits from other businesses.

Table 1: Treatment Effects on Business Outcomes - Extensive

	Business Creation		Main Business				Other Businesses		All Businesses			
	Own a Business		>0 Sales		>0 Profits		>0 Profits		>0 Profits		Extensive Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SBS Implemented	0.142***	0.147***	0.080*	0.080**	0.127***	0.133***	0.065*	0.071**	0.128***	0.136***	0.254***	0.268***
	(0.039)	(0.038)	(0.041)	(0.040)	(0.041)	(0.040)	(0.035)	(0.034)	(0.040)	(0.039)	(0.078)	(0.078)
	[0.010]	[0.010]	[0.079]	[0.050]	[0.010]	[0.010]	[0.079]	[0.050]	[0.010]	[0.010]	[0.010]	[0.010]
GCP Implemented	0.139***	0.138***	0.076	0.076	0.131***	0.131***	0.009	0.008	0.108**	0.109**	0.165*	0.164*
	(0.044)	(0.044)	(0.053)	(0.052)	(0.047)	(0.047)	(0.035)	(0.035)	(0.046)	(0.046)	(0.095)	(0.094)
	[0.010]	[0.040]	[0.366]	[0.366]	[0.030]	[0.050]	[0.782]	[0.881]	[0.069]	[0.089]	[0.208]	[0.277]
Observations	936	1226	937	1227	937	1227	937	1227	937	1227	937	1227
SBS = GCP	0.965	0.880	0.953	0.942	0.949	0.981	0.252	0.197	0.746	0.642	0.471	0.393
Control Mean - Main RCT	0.566	0.566	0.420	0.420	0.486	0.486	0.179	0.179	0.521	0.521	-0.042	-0.042
Control Mean - GCPs	0.619	0.619	0.392	0.392	0.526	0.526	0.103	0.103	0.541	0.541	-0.148	-0.148
Includes Classroom M.	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R^2	0.247	0.233	0.146	0.159	0.201	0.196	0.122	0.105	0.223	0.209	0.244	0.236

Note: Coefficients are ANCOVA estimates that control for the outcome at baseline, the respondent's location, marital status, household size, number of children, age, gender, and level of education, as well as the interaction of these controls with a dummy equal to 1 if the participant was part of the program implemented by a GCP, and 0 otherwise. We record revenues and profits for women without a business as zero to preserve the balance from randomization. Columns (1) and (2) present the reported dummy variable equal to 1 if the participant reports owning a business, columns (3) and (4) presents the reported a dummy variable equal to 1 if the sales for the 3 days prior the survey are greater than 0, columns (5)–(10) present a dummy variable equal to 1 if the self-reported profits for the last month are greater than 0, for either the main business, other businesses or all businesses. Columns (11) and (12) present a summary index for all five outcomes in the table. Odd columns restrict the sample to only OnSite Mentoring and control groups and even columns include Classroom Mentoring and control by an indicator. We report White robust standard errors in parentheses. Romano-Wolf multiple hypothesis test q-values are presented in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2 shows that the impacts of the program on the intensive margin of business out-comes are similar. The program increases new business creation under both implementation modalities. Participants in SBS locations generate approximately 0.2 new businesses over the course of the program compared to the control group, while those in GCP locations generate 0.14 new businesses. Our estimates on sales and profit levels are noisy, but again we do not find significant differences in outcomes between implementers. Qualitatively, it seems that participants in GCP locations focus their efforts on generating higher profits in their main business rather than increasing profits in other businesses. However, the increase in profits from all businesses are proportionally similar between implementers at 13%–16% of respective control mean profits.

Table 2: Treatment Effects on Business Outcomes - Intensive

| | Business Creation | | Main Business | | | | Other Businesses | | All Businesses | | | |
| | No. Businesses | | Sales | | Profits | | Profits | | Profits | | Intensive Index | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SBS Implemented	0.204***	0.227***	4437.75	5001.13	1890.37	3478.19	8380.76	8762.98	13596.69	14665.42	0.07	0.10
	(0.062)	(0.062)	(6052.86)	(5904.93)	(12128.19)	(11788.23)	(5628.40)	(5549.73)	(13344.55)	(13106.68)	(0.09)	(0.09)
	[0.040]	[0.010]	[0.832]	[0.584]	[0.911]	[0.802]	[0.446]	[0.366]	[0.653]	[0.574]	[0.832]	[0.574]
GCP Implemented	0.144**	0.144**	-2773.06	-2742.77	10838.12	10846.94	-3762.82	-3851.32	9343.88	9321.90	0.02	0.02
	(0.071)	(0.070)	(4913.38)	(4873.30)	(12046.13)	(11925.35)	(5068.05)	(5044.03)	(13810.40)	(13692.72)	(0.07)	(0.07)
	[0.267]	[0.248]	[0.891]	[0.891]	[0.851]	[0.851]	[0.891]	[0.891]	[0.891]	[0.891]	[0.891]	[0.891]
Observations	937	1226	912	1197	903	1182	937	1227	937	1227	937	1227
SBS = GCP	0.524	0.373	0.353	0.311	0.603	0.661	0.109	0.093	0.825	0.778	0.642	0.443
Control Mean - Main RCT	0.833	0.833	36484.127	36484.127	69693.200	69693.200	18589.105	18589.105	87084.436	87084.436	-0.044	-0.044
Control Mean - GCPs	0.814	0.814	24163.351	24163.351	58634.737	58634.737	13069.072	13069.072	70701.031	70701.031	-0.092	-0.092
Includes Classroom M.	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R^2	0.296	0.283	0.229	0.230	0.160	0.169	0.116	0.097	0.191	0.188	0.121	0.089

Note: We winsorize all sales and profit measures at the 99th percentile. Coefficients are ANCOVA estimates that control for the outcome at baseline, the respondent's location, marital status, number of children, household size, number of children, age, gender, and level of education, as well as the interaction of these controls with a dummy equal to 1 if the participant was part of the program implemented by a GCP, and 0 otherwise. We record revenues and profits for women without a business as zero to preserve the balance from randomization Columns (1) and (2) present the count of the number of businesses the respondent reports operating, including her main business and all other businesses, columns (3) and (4) presents the reported sales for the 3 days prior the survey, columns (5)–(10) present the self-reported profits for the last month for either the main business, other businesses, or all businesses. All values in columns (3)–(10) are values in UGX. Columns (11) and (12) present a summary index of all five outcomes presented in this table. Odd columns restrict the sample to only OnSite Mentoring and control groups and even columns include Classroom Mentoring and control by an indicator. We report White robust standard errors in parentheses. Romano-Wolf multiple hypothesis test q-values are presented in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Although business results are similar between implementers, are the underlying mechanisms the same? Our results in Table 3 show some differences in hard skills acquisition between the two implementers. Specifically, while both SBS and GCP coaches improve hard skills as measured by the summary index in columns (7) and (8), SBS coaches' impact seems to be slightly higher. The impact of SBS coaches on business tracking, price management, and goal setting is positive and statistically significant, with improvements ranging from 30%–65% of control means. When the program is implemented by GCP coaches, price management practices improve at similar rates as SBS coaches but business tracking and goal setting show qualitatively small effects, significantly so for goal setting.

Table 3: Treatment Effects on Hard Skill Outcomes

	Tracking		Price Mgmt.		Goal Setting		Hard skills Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SBS Implemented	0.320**	0.340**	0.283*	0.314*	0.336**	0.422***	0.300***	0.326***
	(0.149)	(0.143)	(0.166)	(0.161)	(0.154)	(0.147)	(0.086)	(0.084)
	[0.089]	[0.040]	[0.158]	[0.040]	[0.089]	[0.030]	[0.010]	[0.010]
GCP Implemented	0.016	0.023	0.263	0.263	-0.050	-0.049	0.290***	0.290***
	(0.175)	(0.172)	(0.192)	(0.190)	(0.163)	(0.160)	(0.093)	(0.092)
	[0.941]	[0.960]	[0.446]	[0.436]	[0.941]	[0.960]	[0.040]	[0.030]
Observations	488	637	469	616	405	528	933	1222
SBS = GCP	0.186	0.157	0.935	0.839	0.087	0.030	0.941	0.772
Control Mean - Main RCT	0.957	0.957	1.012	1.012	0.643	0.643	-0.060	-0.060
Control Mean - GCPs	0.886	0.886	1.261	1.261	0.529	0.529	-0.056	-0.056
Includes Classroom M.	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R²	0.122	0.123	0.150	0.139	0.102	0.127	0.228	0.223

Note: Coefficients are ANCOVA estimates that control for the outcome at baseline, the respondent's location, marital status, household size, number of children, age, gender, and level of education, as well as the interaction of these controls with a dummy equal to 1 if the participant was part of the program implemented by a GCP, and 0 otherwise. Columns (1) and (2) present the tracking score that combines multiple questions about record keeping for the business with a maximum value of 3. Columns (3) and (4) present the price management score combines multiple questions about setting prices, running promotions, comparing prices with competitors, and negotiating for better prices with suppliers with a maximum value of 4. Columns (5) and (6) present the goal setting score that combines multiple questions about setting goals for the business over various time horizons with a maximum value of 3. Columns (7) and (8) present a summary index of the three scores presented before. Odd columns restrict the sample to only OnSite Mentoring and control groups and even columns include Classroom Mentoring and control by an indicator. We report White robust standard errors in parentheses. Romano-Wolf multiple hypothesis test q-values are presented in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We observe a slightly different pattern for soft skills. Although the program qualitatively increases soft skills for most outcomes among both implementers, treatment effects are relatively small and statistically insignificant. The one exception is that GCP coaches significantly improve the dimension of locus of control that relates a respondent's belief about whether "powerful others" affect their life outcomes. Qualitatively, the implementers appear to affect different types of soft skills, with expansion partners showing larger effects on self-efficacy, grit, and the "powerful others" and "chance" dimensions of locus of control and SBS coaches showing larger effects on the "internality" dimension of the locus of control. However, there are no statistically significant differences in the effectiveness of SBS versus GCP coaches when examining the index of soft skills and the magnitude of the effects on the index are very similar between implementers.

Table 4: Treatment Effects on Soft Skills Outcomes

| | Self-Efficacy | | Grit | | Locus of Control | | | | | | Soft Skills Index | |
| | | | | | Internal | | PO | | Chance | | | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SBS Implemented	-0.003	0.058	0.229	0.217	0.247	0.254	-0.035	0.004	0.349	0.399	0.102	0.110
	(0.567)	(0.558)	(0.507)	(0.496)	(0.204)	(0.203)	(0.418)	(0.405)	(0.354)	(0.345)	(0.083)	(0.081)
	[1.000]	[1.000]	[0.941]	[0.931]	[0.713]	[0.594]	[0.980]	[1.000]	[0.733]	[0.594]	[0.703]	[0.535]
GCP Implemented	0.356	0.349	0.405	0.405	-0.019	-0.016	0.927*	0.927**	0.548	0.547	0.120	0.120
	(0.580)	(0.574)	(0.456)	(0.452)	(0.278)	(0.276)	(0.473)	(0.469)	(0.392)	(0.389)	(0.080)	(0.079)
	[0.812]	[0.782]	[0.743]	[0.683]	[0.950]	[0.980]	[0.297]	[0.248]	[0.554]	[0.416]	[0.545]	[0.386]
Observations	934	1222	934	1223	934	1222	934	1222	935	1223	935	1224
SBS = GCP	0.659	0.716	0.797	0.779	0.439	0.431	0.128	0.137	0.706	0.776	0.878	0.932
Control Mean - Main RCT	38.605	38.605	29.488	29.488	15.836	15.836	-12.914	-12.914	-14.645	-14.645	-0.075	-0.075
Control Mean - GCPs	40.376	40.376	31.577	31.577	16.335	16.335	-13.814	-13.814	-13.753	-13.753	0.179	0.179
Includes Classroom M.	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R^2	0.136	0.122	0.125	0.120	0.074	0.061	0.140	0.149	0.123	0.127	0.181	0.174

Note: Coefficients are ANCOVA estimates that control for the outcome at baseline, the respondent's location, marital status, household size, number of children, age, gender, and level of education, as well as the interaction of these controls with a dummy equal to 1 if the participant was part of the program implemented by a GCP, and 0 otherwise. Columns (1) and (2) present the generalized self-efficacy measure following Schwarzer and Jerusalem (1995), columns (3) and (4) presents our measures of grit following Duckworth, Peterson, et al. (2007) and Duckworth and Quinn (2009), columns (5)–(10) present different locus of control measures. Internal, PO, and Chance is the dimension of the locus of control score. Positive values for Internal, PO and Chance provide evidence of improvements in locus of control measures independent of type. Columns (11) and (12) present a summary index of the scores presented before. Odd columns restrict the sample to only OnSite Mentoring and control groups and even columns include Classroom Mentoring and control by an indicator. We report White robust standard errors in parentheses. Romano-Wolf multiple hypothesis test q-values are presented in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Taken together, results on hard and soft skill acquisition between the two RCTs suggest that the mechanisms underlying overall improvements in business creation and performance may differ between SBS coaches and expansion partners. Importantly, control means across our measures of hard and soft skills are similar between SBS and GCP locations, suggesting that these differences are not driven by dramatically different pre-existing skill levels. Although overall impacts on hard and soft skills are similar between implementers, qualitatively we observe larger impacts on most soft skills outcomes for expansion partners and larger impacts on hard skills from SBS coaches.

5.1 Cost-Benefit

We obtained detailed cost estimates from SBS on the costs of training partner organizations, running the program themselves, and the costs partner organizations report from adding the SBS curriculum to their program. The cost of training one partner coach is USD 5,000.[12] We assume that the cost is the same for an SBS coach, although in reality the cost is likely lower since it does not include travel and accommodation. The highest reported estimate of implementation costs among partner organizations in Uganda is US $95 per participant, while the cost when SBS runs the program itself is US $230 per participant. These differences arise due to costs associated with identifying new communities to serve, higher recruitment costs, and the need to find and rent space to implement the program, all of which are costs that SBS faces but partner organizations do not, or face at lower levels. Finally, SBS estimates that each partner organization will coach around 150 participants. Assuming that each SBS coach implements SBS in 6 communities per year with 50 participants, on average, per community, and that the average SBS coach works for around 5 years, SBS coaches train around 900 participants each.

Putting these estimates together, the per-participant cost of implementing the SBS curriculum averages to US $236 for SBS and US $128 for partner organizations. The difference

---

[12]Note that the program is currently subsidized in the sense that expansion partners are asked to pay at most US $4,900 for *two* coaches to be trained even though the total cost to train two coaches is US $10,000.

is driven by the substantially lower recruitment and implementation costs for partner organizations. Even if we assume that each SBS coach trains 2000 women over her career and costs only USD 1,000 to train, costs are still 80% higher for SBS than they are for partner organizations. The lower costs associated with running SBS as part of existing programs dramatically improves the cost-effectiveness of the intervention.

6 Conclusion

The results of the two studies we conducted indicate that expansion partners are equally effective at implementing the SBS curriculum, and can do so at a substantial cost savings. Our results provide a first proof of concept for a decentralized model of scaling skills-based programs through existing NGOs.

Our study carries limitations. First, we cannot attribute impacts for either SBS or expansion partner coaches to specific mechanisms. For instance, it is not clear whether expansion partners are slightly worse at teaching the curriculum but better at recruiting the participants who will benefit the most from the program, or whether they are more effective at the one-on-one mentoring but less effective at delivering the classroom modules. Second, we cannot observe what activities expansion partners may have undertaken in place of SBS. If we view local NGOs as socially motivated organizations who are capable of identifying and implementing the interventions with the highest social returns, then we can interpret their choice to implement SBS as optimal. However, it may be difficult for such organizations to identify the programs with the highest social returns. Better understanding how local NGOs decide on which programs to implement is central to fully accounting for the impacts of the program.

Given that our work merely presents an initial proof of concept, there are a multitude of open questions remaining for future research. First, it will be important to continue testing and expanding the external validity of our results. The program scaled well in Uganda, where

it was developed, but SBS currently has expansion partners throughout Africa as well as in Asia and South America. Another dimension of external validity is to study the scalability of other skills-based programs. Are the effects we observe unique to the manner SBS teaches its curriculum to expansion partners? Our results also raise questions about the most effective way to scale programs in partnership with governments. If government-led programs scale through decentralized local governments rather than national-level initiatives, can they better preserve impact? Building and testing different decentralized paths to scale will illustrate the most effective elements of the scaling model we study and provide answers to the pressing issue of how to scale effective programs for economic development.

References

Angelucci, Manuela, Dean Karlan, and Jonathan Zinman (2015). "Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco". *American Economic Journal: Applied Economics* 7, 151–182.

Araujo, M. Caridad, Marta Rubio-Codina, and Norbert Schady (2021). "70 to 700 to 70,000: Lessons from the Jamaica Experiment". *IDB Working Paper Series*.

Attanasio, Orazio et al. (2015). "The Impacts of Microfinance: Evidence from Joint-Liability Lending in Mongolia". *American Economic Journal: Applied Economics* 7, 90–122.

Augsburg, Britta et al. (2015). "The Impacts of Microcredit: Evidence from Bosnia and Herzegovina". *American Economic Journal: Applied Economics* 7, 183–203.

Bando, Rosangela, Emma Naslund-Hadley, and Paul Gertler (2019). "Effect of Inquiry and Problem Based Pedagogy on Learning: Evidence from 10 Field Experiments in Four Countries". *NBER Working Paper*.

Banerjee, Abhijit, Rukmini Banerji, et al. (2017). "From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application". *Journal of Economic Perspectives* 31, 73–102.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, et al. (2015). "The Miracle of Microfinance? Evidence from a Randomized Evaluation". *American Economic Journal: Applied Economics* 7, 22–53.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, et al. (2015). "A multifaceted program causes lasting progress for the very poor: Evidence from six countries". *Science* 348.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman (2015). "Six Randomized Evaluations of Microcredit: Introduction and Further Steps". *American Economic Journal: Applied Economics* 7, 1–21.

Crépon, Bruno et al. (2015). "Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Experiment in Morocco". *American Economic Journal: Applied Economics* 7, 123–150.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar (2014). "Keeping It Simple: Financial Literacy and Rules of Thumb". *American Economic Journal: Applied Economics* 6, 1–31.

Duckworth, Angela L., Christopher Peterson, et al. (2007). "Grit: perseverance and passion for long-term goals". *Journal of Personality and Social Psychology* 92, 1087–1101.

Duckworth, Angela Lee and Patrick D. Quinn (2009). "Development and validation of the short grit scale". *Journal of Personality Assessment* 91, 166–174.

Duflo, Esther (2004). "Scaling Up and Evaluation". *Annual World Bank Conference on Development Economics*.

Lang, Megan and Julia Seither (2022). "The Economics of Women's Entrepreneurs: Evidence from Building Skills in Uganda". Working Paper.

— (2024). "Building Women's Skills for Economic Inclusion and Resilience". Policy Research Working Paper No. 10980, World Bank.

Romano, Joseph P. and Michael Wolf (2005). "Stepwise Multiple Testing as Formalized Data Snooping". *Econometrica* 4, 1237–1282.

Schwarzer, R. and M. Jerusalem (1995). "Generalized self-efficacy scale". *Measures in Health Psychology: a user's portfolio*.

Tarozzi, Alessandro, Jaikishan Desai, and Kristin Johnson (2015). "The Impacts of Microcredit: Evidence from Ethiopia". *American Economic Journal: Applied Economics* 7, 54–89.

Al-Ubaydli, Omar, Min Sok Lee, et al. (2021). "The Science of Using Science". *The Scale-Up Effect in Early Childhood and Public Policy*. Ed. by John List, Dana Siskind, and Lauren Supplee. Taylor & Francis, 104–125.

Al-Ubaydli, Omar, John A. List, and Dana Suskind (2019). "The Science of Using Science: Towards an Understanding of the Threats to Scaling Experiments". *NBER Working Paper*.

Vivalt, Eva (2020). "How Much Can We Generalize from Impact Evaluations". *Journal of The European Economic Association* 18, 797–816.

A   Appendix

A.1   Figures

Month 0	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6
- Mobilization - Orientation (aspirations intervention)	- "Getting out of your comfort zone" - Identifying business opportunities - Finding capital and starting small	- Bookkeeping and record keeping - Market research	- Business planning	- Growing your customer base - Money management	- No modules (implement business plans)	- Graduation Ceremony

Note: This figure is taken from Lang and Seither (2022).

Figure A1: SBS Module Content

A.2   Tables

Table A1: Descriptive Statistics - Main RCT (Only Intensive Mentoring) vs. GCP Participants

Variable	(1) Main-Control N	Mean/(SD)	(2) Main-Treatment N	Mean/(SD)	(3) GCP-Control N	Mean/(SD)	(4) GCP-Treatment N	Mean/(SD)	N	F-test for balance across all groups F-stat/P-value
Male	285	0.000 (0.000)	339	0.000 (0.000)	203	0.084 (0.278)	191	0.173 (0.379)	1018	36.811*** 0.000
Age	285	37.870 (12.919)	339	37.454 (12.274)	203	35.670 (13.179)	191	37.267 (12.686)	1018	1.289 0.277
Married	285	0.618 (0.487)	339	0.622 (0.485)	203	0.611 (0.489)	191	0.696 (0.461)	1018	1.394 0.243
Divorced	285	0.182 (0.387)	339	0.201 (0.401)	203	0.133 (0.340)	191	0.120 (0.326)	1018	2.645** 0.048
Single	285	0.084 (0.278)	339	0.071 (0.257)	203	0.163 (0.370)	191	0.094 (0.293)	1018	4.419*** 0.004
Widowed	285	0.116 (0.321)	339	0.106 (0.309)	203	0.094 (0.292)	191	0.089 (0.285)	1018	0.379 0.768
Primary Ed.	285	0.488 (0.501)	339	0.522 (0.500)	203	0.596 (0.492)	191	0.644 (0.480)	1018	4.743*** 0.003
Secondary Ed.	285	0.902 (0.298)	339	0.950 (0.219)	203	0.936 (0.245)	191	0.948 (0.223)	1018	2.237* 0.082
Father Primary Ed.	285	0.765 (0.425)	339	0.785 (0.412)	203	0.813 (0.391)	191	0.869 (0.338)	1018	2.886** 0.035
Father Secondary Ed.	285	0.933 (0.250)	339	0.926 (0.262)	203	0.980 (0.139)	191	0.984 (0.125)	1018	4.821*** 0.002
Mother Primary Ed.	285	0.828 (0.378)	339	0.885 (0.320)	203	0.901 (0.299)	191	0.901 (0.300)	1018	2.845** 0.037
Mother Secondary Ed.	285	0.954 (0.209)	339	0.973 (0.161)	203	0.985 (0.121)	191	0.995 (0.072)	1018	2.926** 0.033
Employed	285	0.537 (0.500)	339	0.496 (0.501)	203	0.611 (0.489)	191	0.592 (0.493)	1018	2.883** 0.035
HH Size	285	4.225 (2.488)	339	4.230 (2.620)	203	4.931 (2.678)	191	5.021 (2.973)	1018	6.359*** 0.000
Minors in HH	285	2.926 (2.101)	339	2.991 (2.099)	203	3.419 (2.026)	191	3.586 (2.304)	1018	5.420*** 0.001
Own a Business	285	0.551 (0.498)	339	0.490 (0.501)	203	0.517 (0.501)	191	0.466 (0.500)	1018	1.324 0.265

Notes: Mean baseline covariates by treatment group. Standard deviations are in parentheses. Column 5 reports p-values associated with F-tests of joint equality between the three groups. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A2: Attrition

	(1) At Exit	(2) At Exit
SBS OnSite M.	0.039	0.046*
	(0.021)	(0.021)
GCP OnSite M.	-0.050**	-0.042*
	(0.017)	(0.018)
Male		-0.029
		(0.031)
Age		-0.000
		(0.001)
Married		0.058*
		(0.024)
Divorced		0.033
		(0.028)
Single		0.085
		(0.043)
Primary Ed.		-0.004
		(0.019)
Secondary Ed.		-0.031
		(0.043)
Father Primary Ed.		0.013
		(0.026)
Father Secondary Ed.		0.030
		(0.046)
Mother Primary Ed.		0.014
		(0.031)
Mother Secondary Ed.		0.005
		(0.063)
Employed		0.040
		(0.022)
HH Size		-0.001
		(0.009)
Minors in HH		-0.010
		(0.011)
Daily Expenditure HH		0.000
		(0.000)
Savings		-0.000
		(0.000)
Remittances		-0.000
		(0.000)
Own a Business		-0.064**
		(0.022)
Sales - MB		-0.000
		(0.000)
Profits - MB		-0.000
		(0.000)
Profits - OB		-0.000
		(0.000)
Locus - Internal		0.006
		(0.003)
Locus - PO		-0.001
		(0.002)
Locus - Chance		0.002
		(0.003)
Observations	1018	968

Note: For the marital status, the omitted dummy is the Widowed status. The following are abbreviations used in the table: MB stands for Main Business, OB stands for Other Businesses, HH stands for Household. For the definitions and descriptions of all other variables, see Appendix.