Journal of Development Economics

Improving Childcare Quality Through Social Franchising

--Manuscript Draft--

Manuscript Number:	DEVEC-D-24-01906R2
Article Type:	Registered Report Stage 1: Proposal
Section/Category:	Health, Education, gender, poverty
Keywords:	childcare; daycare; quality; profits; microenterprise; Labor supply; child development
Corresponding Author:	Anne Fitzpatrick The Ohio State University Columbus, OH UNITED STATES
First Author:	Anne Fitzpatrick
Order of Authors:	Anne Fitzpatrick
	Emily Beam, Ph.D.
	Maira Reimao, Ph.D.
Abstract:	Childcare quality is important for children's development and parental labor market outcomes. However, in low-income settings, private childcare often sits at an equilibrium of low prices and low quality, where parents have limited willingness to pay and profitability is low. We partner with a social enterprise dedicated to improving childcare quality through training, facility improvements, and continuous in-kind support. We randomize entry into 51 informal settlements across 11 counties in Kenya. With 978 providers, we test whether childcare quality improves, whether providers revenues change in response to quality, and whether daycare enrollment changes 12, and 24 months after entry. We combine three rounds of firm data with data from 2,820 households to also examine the effect of improving childcare quality on household labor force participation and children's development. This study informs whether a social-franchising model can improve childcare quality and improve market functioning in settings with limited regulation.
Response to Reviewers:	

Improving Childcare Quality through Social Franchising: A Pre-Results Review Report Proposal

April 24, 2025

Abstract

Childcare quality is important for children's development and parental labor market outcomes. However, in low-income settings, private childcare often sits at an equilibrium of low prices and low quality, where parents have limited willingness to pay and profitability is low. We partner with a social enterprise dedicated to improving childcare quality through training, facility improvements, and continuous in-kind support. We randomize entry into 51 informal settlements across 11 counties in Kenya. With 978 providers, we test whether childcare quality improves, whether providers revenues change in response to quality, and whether daycare enrollment changes 12, and 24 months after entry. We combine three rounds of firm data with data from 2,820 households to also examine the effect of improving childcare quality on household labor force participation and children's development. This study informs whether a social-franchising model can improve childcare quality and improve market functioning in settings with limited regulation.

Keywords: childcare, daycare, quality, profits, microenterprise, labor supply, child development

JEL Codes: J13, D1, D4, O12

Study Pre-Registration: This study is registered as AEARCTR-0011747, "Improving Childcare Quality through Social Franchising"

Publication timeline

If accepted to the JDE pre-results review, our results will be published after the collection and cleaning of the endline data. This date is tentatively scheduled for October 2026.

1 Introduction

Access to high-quality paid childcare ("daycare") is central to advancing women's economic and social agency while also promoting child development.[1] However, high-quality daycare is out of reach for many parents in both high- and low-income settings (Neuman and Powers, 2021). Parents may be unable to pay higher prices for quality care or be unwilling to do so given competing demands when quality is not fully observable (Devercelli and Beaton-Day, 2020). In Kenya, where our study takes place, societal norms to look after children may further pressure providers to provide care despite late or missing payments from families, and these structural supply-side challenges compound into notoriously low profits in the industry as a whole (Blau, 2001). As a result, low quality may result in an equilibrium where there are few daycare firms and the daycare firms that exist are unprofitable. As a result, both children and their parents also miss out on the benefits of high-quality childcare, ranging from higher household incomes, lower maternal stress, and improved child development.

In this paper, we analyze whether assisting daycare providers improve their quality also improves revenues, and we document downstream effects on outcomes for households and children. We partner with a Kenya-based social enterprise that works with owners of existing private daycare facilities in informal settlements throughout Kenya and provides training, mentorship, in-kind support, and quality assurance. Firms that complete the program and meet data quality standards are invited to become franchise owners, for which they pay a monthly fee scaled to their firm size. Franchisee status indicates to outsiders that the center meets a minimum quality standard set by the partner organization. Firms that opt to become a franchisee also receive a one-time capital improvement grant worth approximately $200, plus regular mentoring and in-kind feeding support for the children they serve.

We conduct a clustered randomized evaluation in 51 communities spread across 11 counties of Kenya to evaluate the impacts of this social franchising model. Specifically, we randomize the entry of our partner organization into the communities and assess the impacts on both the supply and demand for daycare. This study has high external validity: 51% of urban residents in Kenya live in informal settlements (slums); globally, 24% of urban residents live in these types of areas (World Bank, 2020).

This paper will make three contributions to the academic and policy literature on childcare and microentrepreneurship.

First, we contribute to the childcare literature by identifying the impacts of high-quality childcare. The existing literature primarily identifies how increased access to daycare changes child development and/or maternal labor supply. These results suggest that without daycare, maternal earnings are lower (Delecourt and Fitzpatrick, 2021; Delecourt et al., 2022), and daycare can potentially be transformative for families. For example, improving access by providing subsidies or creating new centers improves maternal (and paternal) earnings (Bjorvatn et al., 2022; Ranganathan

[1] Throughout this paper, we refer to market-provided care as "daycare," rather than "childcare," which we think of as also including unpaid care provided by household members.

and Pedulla, 2018; Cascio, 2009; Donald et al., 2023). In addition to earnings and financial outcomes, Ajayi et al. (2022) finds that access to community-based daycare can also lead to small reductions in depression and self-reported unhappiness among mothers. A recent study by Clark et al. (2019) from the same setting as ours found that a free voucher for childcare increased household earnings by 8% and encouraged mothers to switch into occupations for which it is difficult to bring children.

Investments in early childhood education in high-income contexts have been shown to increase a variety of long-term outcomes ranging from the likelihood of having a bank account to voting to completing high school (Heckman and Karapakula, 2019; Cohodes and Feigenbaum, 2021; Gray-Lobe et al., 2021; Bartik, 2022). Children's development is also improved following improved access to childcare (Bjorvatn et al., 2022; Donald et al., 2023), and improving public sector quality has been shown to have substantial benefits for children, especially those from disadvantaged households (Andrew et al., 2024). Yet, some studies have found that participation in childcare harms children (Baker et al., 2008) or that the benefits fade out over time (Jakiela et al., 2024), suggesting that the quality of care received is crucial for children to receive and maintain its benefits. By contrast, there is scant attention paid to supply-side considerations that lead to insufficient use of daycare, such as the extremely low quality and low profitability of firms in the market. Analyzing the role of quality on use and sustainability is important to understanding the paradox of insufficient use coupled with high reported maternal demand for additional care options (Hughes et al., 2023). Although the literature has emphasized high child-to-staff ratios, often exceeding 30:1, scant access to basic learning materials, less attention has been paid to basic concerns of insufficient feeding and care practices that may stunt development or even result in toxic levels of child stress (Devercelli and Beaton-Day, 2020). We fill this gap by analyzing whether improving quality among firms improves revenues in a resource-constrained environment and looking at overall impacts on the quantity of children served in the market, with downstream effects on households and children.

Second, we provide new evidence to microenterprise literature on how quality improvements affect the profitability of microenterprises and overall market dynamics. A large literature has focused on finding policies and interventions that increase microenterprise profits, firm size, and business practices, such as capital grants (De Mel et al., 2013; Fiala, 2013), mentorship (Brooks et al., 2018; McKenzie and Puerto, 2021), business training (Bakhtiar et al., 2022; McKenzie et al., 2021), or consulting (Bruhn et al., 2018; Anderson and McKenzie, 2020).[2] There is comparatively less work on programs to improve the *quality* of goods or services in the private sector in low-income settings. We fill that gap by collecting a rich set of data measuring the quality of daycare firms and testing a way to improve these outcomes.

The design of our study also has several novel characteristics that add to its external validity and impact. The literature on microentrepreneurship has largely relied on randomization at the individual level and therefore captures data on a relatively small number of firms. By contrast, our study is large, covering 51 different markets, allowing us to make a unique contribution by

[2]See Jayachandran (2021) or Quinn and Woodruff (2019) for a more extensive summary of the literature.

measuring the causal market-level and spillover effects of individual firm quality improvements. This contribution is notable as competitor responses to improved quality competition is theoretically ambiguous (Gaynor and Town, 2011). In particular, non-franchised firms may lower both prices and quality in a "race to the bottom," leading to market segmentation into "high quality" and "low quality" firms. Alternatively, lower-quality firms may exit, reducing overall service availability. Previous empirical work has found that the introduction of a new grocery competitor reduces prices and increases service quality in the retail sector (Busso and Galiani, 2019). However, there is less work on the market response to changes in the quality competitors offer. One exception is McKenzie and Puerto (2021), who improve firm quality via training and mentorship and find little spillover effects onto competitor firms. However, the nature of childcare—an experience good that is purchased more frequently than a typical retail enterprise—may lead to different outcomes, as parents may not even recognize or be able to pay for higher-quality care. Thus, our study may provide novel and rare empirical evidence on how firms adjust operations and prices to quality improvements and how their competitors respond.

Third, our results inform policymakers worldwide about the causal impact market-based interventions to improve childcare quality. Understanding whether improvements in quality increase or decrease firm profitability and increase or decrease overall use of daycare is crucial for advancing knowledge on whether improving daycare quality may have unintended consequences. Research on supply-side responses to quality improvements is surprisingly scant. The bulk of the rigorous evidence comes from high-income settings, where, for example, tightening state childcare regulations in the U.S. has been shown to reduce the number of center-based providers—especially in lower-income markets (Hotz and Xiao, 2011). However, this intervention is not direct regulation, but rather quality improvements that parents may or may not recognize. Previous work in the context of primary education in rural Pakistan has emphasized that education quality is unobservable to parents, but that parental information about quality ultimately results in both increases in enrollment and further increases in quality (Andrabi et al., 2017). Similarly, several studies have found that parents are willing to pay for higher education quality (see, for example: Carneiro et al. (2024)). However, whether those insights apply to daycare providers in urban Kenya are uncertain. On the one hand, daycare quality is likely less observable to parents in this market as children may have limited verbal capacity, and parents may not notice improvements in quality. On the other hand, parents in our urban context may have numerous options for daycare, and thus may be particularly responsive to quality improvements. This study will therefore provide insight on whether improving quality helps to regulate daycare quality, and whether it results in spillover effects or firm closures. Altogether, these contributions are substantial and important, regardless of whether we find expected results or null results. Furthermore, academic research on quality competition is nascent and often not focused on this key industry. We fill this gap and provide new data points for models of quality competition and how labor force participation and health are ultimately affected.

In this document, we pre-specify and outline our hypotheses, data collection, and analysis plan to provide estimates on these important contributions to both the academic and policy literature.

For providers, we examine the impact on daycare quality, enrollment, and revenue as our primary outcome measures. For providers, we also examine impacts on profits and provider agency and pride in the profession as a second step (contingent on observing an effect on primary outcomes, following our theory of change). For households, we consider as secondary outcomes impacts on labor force participation and employment of caretakers, along with select measures of child development. We also leverage our research design to look at market-level impacts, specifically firm entry and exit, and subsequent impacts on the market from the emergence of high-quality competitors. We will follow up with firms in early 2025 to collect initial data on revenues and enrollment; then again in mid-2025, approximately 6 months after completion of the training (12 months after entry), to measure short-run program impacts on revenue and quality. We will conduct a second follow-up with firms and households in mid-2026, approximately 18 months after completion of the training (24 months after entry) to measure the longer-term impacts on firm outcomes as well as child development and household employment outcomes.

2 Intervention

We evaluate an existing intervention designed to improve daycare quality in informal settlements in Kenya by establishing a network of vetted and supported childcare centers in these areas. In its normal operations, the partner organization identifies communities that it will enter according to need, safety, and other viability considerations. Together with local health workers, it canvases the chosen communities and invites all eligible daycare providers to a free, half-day workshop in which they introduce their social enterprise model to attendees (owners, principals, or teachers) and explain their program, which involves training, nutritional support, and capital transfers.

After this half-day introduction, interested daycare owners can opt to participate in a three-day training for a fee of 200 KES (USD 1.50).[3] This training includes lessons in effective early education practices and maintaining health and hygiene in daycare centers as well as topics related to business and transforming the self-image of the daycare owners and directors (largely women in this context) from childminders to professional caregivers, as the aim of this program is not only to improve but also professionalize daycare services. Following this short training, those who want to become a franchisee pay an additional fee of 1000 KES and join a three-month mentoring program, during which time they receive regular visits from mentors and can opt to participate in additional childcare professional development sessions. The mentoring and training cover a wide variety of topics including (but not limited to): provider-child interactions, safety and health (WASH), business training and bookkeeping, and teaching and learning materials. For example, providers in the program may learn from their mentor how to make books out of cardboard and other readily available materials for the children in their centers. Daycare providers also learn to reuse items like yogurt cups and bottle tops as low-cost teaching and learning materials.

[3]We use the average exchange rate from March–May 2024, $1 USD = KES 132.76, when converting from Kenyan Shillings to USD (OANDA, 2024).

Following the three-month program, owners whose centers meet quality standards may choose to become a franchisee, which they can advertise to the community. They receive a one-time capital-improvement grant to upgrade the physical characteristics of their establishment to ensure that it is safe and hygienic.[4] This grant is worth approximately US$200, and its use varies according to the needs of each member. Owners also receive a smartphone with the partner's app; continued mentoring and support from a person from within their own community hired to fulfill that role (this individual works to ensure best practices, answer questions, and generally support the daycare owners); and regular delivery of fortified porridge for all participant children. To maintain their franchisee status—which entails this continuous support and the ability to display and use the franchise logo—daycare owners pay a franchise fee of KES 500–1500 per month (USD 4–11.50), depending on the number of children served. This fee is not payable during school breaks, and our implementing partner has noted that the requirement has not been strictly enforced; only approximately half of providers regularly pay fees on time.[5]

Workshop

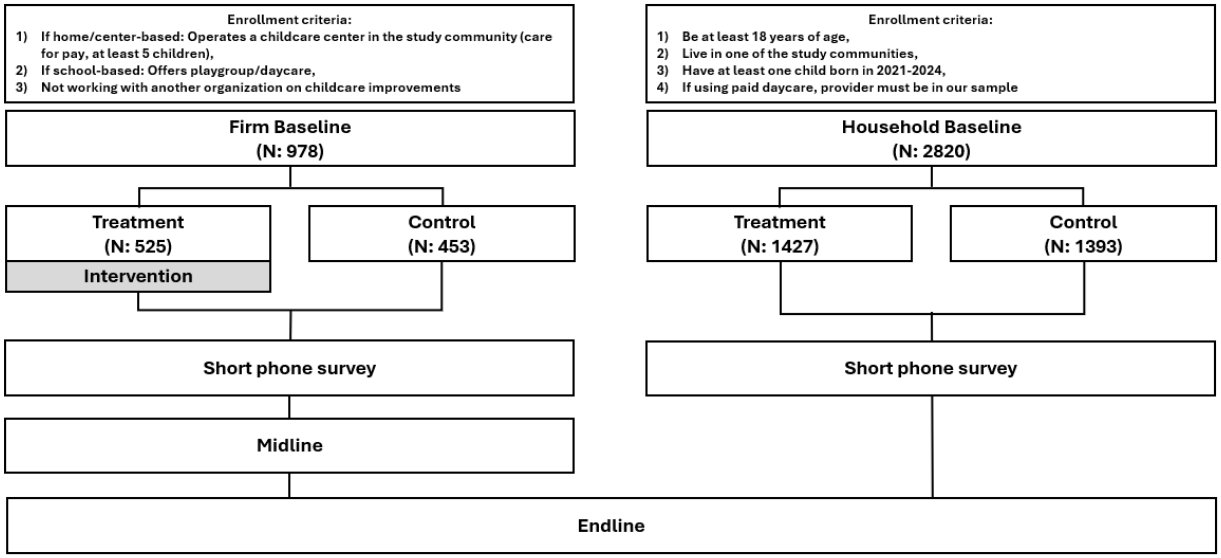
According to our implementation partner, the take-up rate of the social franchise (i.e., participating in the training program and ultimately opting to become a franchisee) among daycare providers in the areas they currently operate is between 30% and 50%, as center owners can choose to join as a franchise or not. To increase statistical power, as well as identify the counterfactual likely-enrollees in the control group, our partner organization agreed to conduct a half-day workshop in both treatment and control groups in May 2024. In control groups, this workshop covered "responsible caregiving and the role of a caregiver," a topic that was selected to be relevant to participants and worth their time but not in sufficient depth in such a short time to lead to longer-term outcomes and because our implementation partner also includes this topic within its regular training program. In treatment communities, our partner delivered its usual half-day workshop, introducing participants to its program, including its mission and values, and offering an overview of the training and mentorship program. In both treatment and control areas, representatives of the partner organization visited all daycare centers and invited them to the workshop. The recruitment procedure was identical in both treatment and control communities and, in both arms, the workshops were delivered inside the communities to all interested center owners or managers, lasted half a day, and were free.

In total, 583 providers (59.6%) in our sample participated in the workshop, and the take-up rate is similar in treatment and control communities: 58.3% in control and 60.8% in treatment. This difference is not statistically significant ($p = 0.499$). Reassuringly, we find no evidence of differences

[4]For example, as part of their capital improvement grant, our partner may facilitate the building of a changing table and create a separate location for diapering. Alternatively, they may erect barriers to keep children away from kitchen areas or open flames used during meals. The exact improvement will vary depending upon the needs and the *status quo* of a particular provider.

[5]In collaboration with our implementing partner, we opted to have them treat the study centers the same as their other member centers for purposes of enforcement, as we felt it would be most useful to learn about the effects of the program as it currently exists.

Figure 1: Study Design Figure



in observable characteristics between workshop participants in treatment and control communities (See Appendix Tables A1 and A2).

3 Evaluation design

3.1 Methodological framework

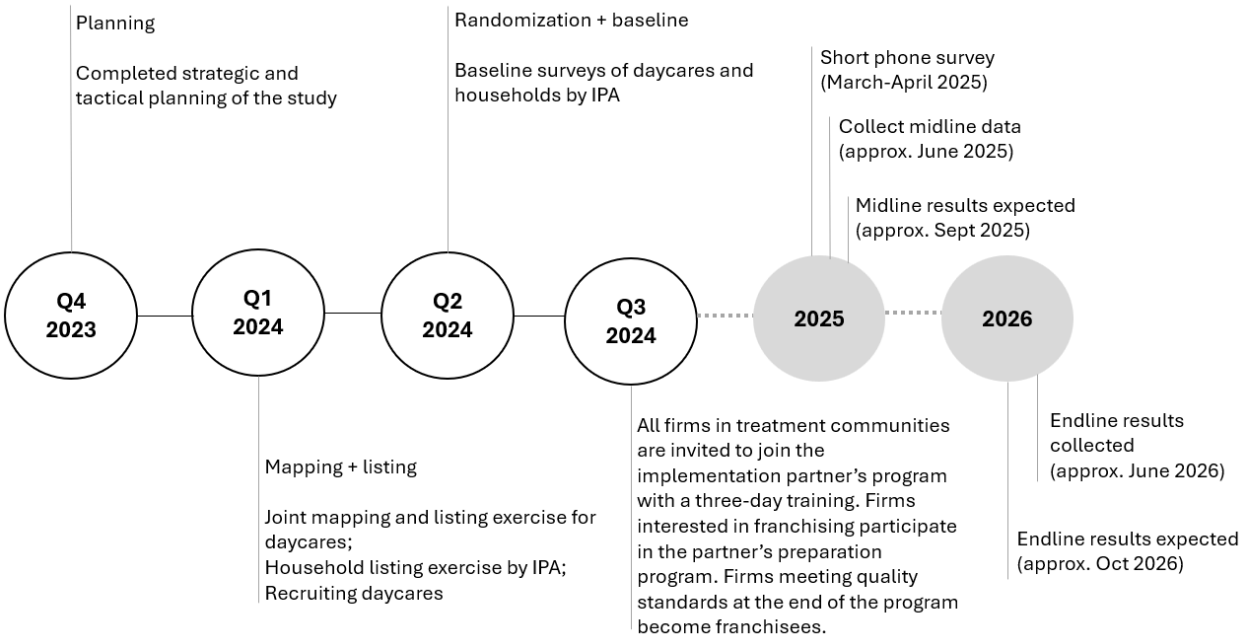
To evaluate the impact of improved childcare quality, we design a clustered randomized controlled trial (RCT), where each cluster is a community. We define "community" in our evaluation as a contiguous geographic area within an informal settlement. The RCT described below includes 51 communities across 11 counties that our partner organization identified in 2023 as communities they would be interested in serving and willing to enter.

For this study, we define a daycare provider as any entity caring for at least five children aged 3 and under and charging a fee for at least some of those children. These include home-, center-, and school-based providers. We present our study design graphically in Figure 1; our study timeline is in Figure 2.

3.2 Theory of change

We hypothesize that due to training, mentorship, a grant that improves the quality of infrastructure, and fortified porridge deliveries, our partner organization will directly improve the quality of care offered by treated providers through improved safety, sanitation, child engagement, and the provision of food. Profit-maximizing providers will then increase prices. Parents will feel more comfortable sending their children to these providers, and enrollment at these facilities will also increase

Figure 2: Study Design Timeline



despite higher prices.[6] As a result, revenues and profits will increase and persist. We hypothesize that part of this response may reflect increased market segmentation. As this is a low-income, highly price-sensitive setting, some parents may be unwilling or unable to pay for the quality improvements. Thus, competitors who have not undergone the training or become franchisees might not improve quality and instead compete on price.

For households, we hypothesize that as a result of the increased availability of high-quality daycare in their area, they will be more likely to enroll their child in daycare or increase their use of daycare. With this, parents will be more likely to work; will work more hours; and will earn more money, partly as a result of switching occupations from casual day labor to more regular jobs. We also hypothesize that child development will improve due to a combination of improved feeding; improved stimulation; and improved parental income. Parents who are priced out of the improved daycares may continue to use the non-franchisee providers.

Because the childcare sector is largely informal, parents may have limited information about how care quality differs between centers and which centers offer high-quality care. Thus, how quickly information regarding improved quality – and that franchising (and a logo) is associated with meeting quality standards – spreads through the market is an open question. We will specifically seek to measure the attributes that households value early on to determine whether households indeed have a willingness to pay for quality, and whether they are aware of any high-quality providers in their area.

[6]The effect on child:caretaker ratios will depend on whether firms are currently operating at capacity or not.

3.3 Hypotheses

In this section, we present and discuss the primary and secondary hypotheses in this study, which span both the childcare and entrepreneurship literature. The survey questions associated with each of these these hypotheses are listed in Appendix Tables A3 and A4.

3.3.1 Primary Hypotheses

The following hypotheses are central to our study and we have designed it to detect changes in their associated outcomes:

1. **HA1. The social franchising model will improve daycare quality**, where quality is defined according to different taxonomies below.

   Daycare quality is a multi-dimensional aspect of substantial importance to the study and its conclusions. We define three sub-indices of daycare quality based on aggregating the set of binary indicators. We will also sum them into one aggregate "Firm quality index."

   We emphasize here that, while we expect that improvements in quality will come through given nature of the invention, these changes are important and essential to child wellbeing given baseline conditions. For example, in our baseline survey, 44% of daycare centers reported not having any toys and 37% not having any books, revealing that a large share of children are in facilities with limited resources for stimulation or supporting cognitive development. We present descriptive statistics of some baseline quality conditions in Appendix Table A5, and highlight that improvements in these outcomes would be materially meaningful.

   - **Hygiene and Safety Index** (Questions H04, H20, and M03). This index encompasses several key dimensions, including the presence of a basin or sink for handwashing, a jug or bottle for handwashing, a first-aid kit or box, potties for toilet training, a changing table or diapering area, and mattresses or mats for napping.
   - **Toys and Manipulatives Index** (Question H05). This index covers the types of toys and manipulatives available to children at the facility.
   - **Child Experience Index** (Questions H01, H02, H03, H04, H07, K07). This index covers various aspects of the child experience in daycare, including caregiver reports of whether children follow a daily schedule, if they sing songs or read to the children. We also include enumerator observations of whether caregivers played with or engaged with children, as well as the presence of a posted schedule or timetable, the availability of an outdoor play area, and the presence of wall decorations.

2. **HA2. The social franchising model will increase provider revenue**.

   Firm revenue (Questions Q1, Q2, and Q3) will be calculated using the number of children in attendance on each day of the week and rates charged (which may vary by length of time; age; payment compliance rate, and additional fees received).

We log and winsorize revenue at the 99th percentile.

3. **HA3. The social franchising model will increase the number of child-days served**.

   The child-days served will be based on asking each firm how many children attended their center in the past 7 days under different attendance schemes (Questions Q1 and Q2).

3.3.2 Secondary Hypotheses

We also consider the following outcomes as a second step, following our theory of change. We note that these are outcomes that are not secondary in importance, but, rather, that we expect to change only if we find an effect on primary outcomes. Our study is not specifically powered to detect these changes.

1. **HB1. The social franchising model will have spillover effects onto competing providers**, as measured by market entry (or exit), as well as the total number of children attending daycare, aggregated to the community level.

2. **HB2. The social franchising model will increase provider profits**.

   We will also measure impact on **profit** but note that this is a secondary outcome because we expect considerable imprecision in costs, particularly for school-based centers, for which it is harder to think of or calculate the marginal cost of just the daycare/ playgroup classrooms.

   We log and winsorize profits at the 99th and 1st percentile, and we log and winsorize costs at the 99th percentile.

3. **HB3. The social franchising model will increase provider agency and pride in profession.** We will create an index of variables related to agency and empowerment. These include whether the respondent thinks that running a daycare is more respected in their community than other common competing sources of female entrepreneurship in this context (selling chapati or styling hair, D03 and D04); questions on self-efficiacy and social standing (D08-D09) as well as a ladder question (D18).

4. **HB4. The social franchising model will increase household usage of childcare**.

   Household's daycare use, extensive margin: Whether a child under age 4 (up to 47 months) ever attended daycare (regardless of type) school in the past 30 days (Question G3).

   Household daycare use, intensive margin: Number of days a child under age 4 attended a daycare center or school in the past 30 days (Question G10).

5. **HB5. The social franchising model will increase maternal employment and household earnings**, as measured by the labor market activities module.

   Primary caregivers' labor supply, extensive margin: This variable is measured by whether the child's primary caregiver has worked for pay in the past 7 days. This could

include (a) working for wages/salary (including domestic and farm work) (Question F1); (b) working on a business/income-generating activity they own or operate (Question F5); and/or (c) helping on a business or income-generating activity that someone else owns or operates (Question F9). The child's primary caregiver is self-designated by the respondent to the household survey at baseline.

Primary caregivers' labor supply, intensive margin: Number of hours child's primary caregiver worked for pay in the past 7 days, summed across the above 3 categories (Questions F3, F7, and F11). We plan to include this variable both (a) imputing zeros for households that are either unemployed or not in the labor force; and (b) conditioning values among households that are employed.

Primary caregivers' earnings over the past 30 days (Question F4 and F8). We calculate the caregiver's total earnings by adding their income from employment and their income from running their own business.

6. **HB6. The social franchising model will improve select measures of child development**, as measured by caregiver-reported benchmarks across several domains.

   Note that these child development measures were not collected at the time of the baseline survey. We will file an amendment to the pre-analysis plan once the post-intervention household survey scheduled for June 2026, which will include caregiver-reported child development questions, is appropriately piloted.

7. **HB7. The social franchising model will improve child nutritional intake** through increased daycare feeding.

   According to baseline survey data, 31% of daycares in this setting do not provide any meal or snack to children (Question F02), and 40% report that, at least once or twice a week, a child might arrive without food and is hungry throughout the day (H14).

3.4 Data and sample

In this section, we outline how we defined a community, the unit of randomization. We next explain how we selected daycare owners to participate in the provider survey and how we selected households to participate in the caregiver survey.

3.4.1 Boundary mapping communities and listing daycares

In 2023, our implementation partner identified 55 potential neighborhoods for entry. However, while the general neighborhoods of informal settlements are known, there are no clear administrative boundaries. Thus, for purposes of this study we conducted an exercise to delineate the boundaries of each "community" with the support of local community health providers, local leaders, and our partner organization. This exercise was done in January 2024 by walking in each neighborhood, using GPS to trace polygon boundaries of the communities. The end result is a set of maps that

delineate each community's boundary in our study; each of these bounded areas is a community for the purposes of this research.

Once the mapping was complete, we canvassed the mapped communities to create a census of all daycare providers in each community that serve children ages 3 and younger.[7] These providers are further distinguished between (1) home- and center-based providers and (2) school-based providers. In the latter group, daycare/ playgroup classrooms are offered within a school that also serves kindergarten and primary school students. Figure 4 provides a sample (jittered) map with the identified providers in a community.

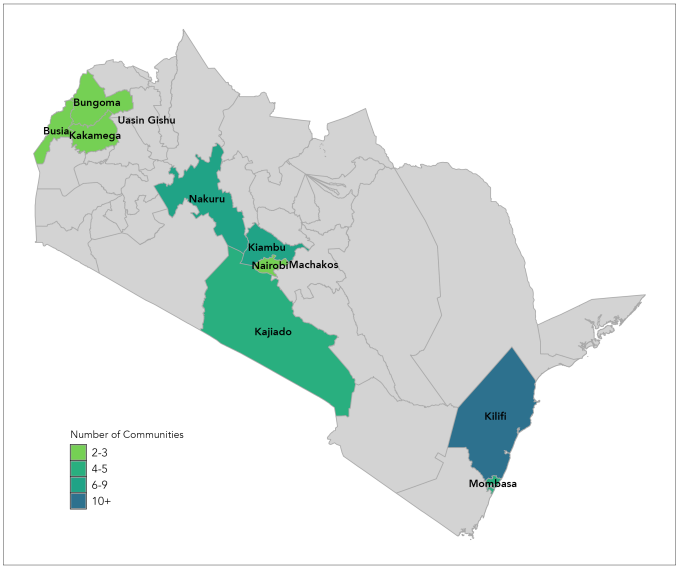
Through this phase, communities were jointly screened by our partner and field staff for our study primarily on the basis of security and enumerator safety. Additionally, our partner organization required approximately 10 daycares to operate in a community in an effort to make the intervention's costs sustainable.[8] Thus, some candidate communities were either excluded from the study or combined with nearby communities to meet the minimum requirement for operations per area. Finally, we ensured that many daycare providers within the community were not already receiving benefits from any other care-focused NGO.[9] The result of this process is 51 communities suitable for randomization. Prior to randomizing at the community level, we conducted baseline surveys of both firms and households, as described in the next section.

[7] As these are informal providers working in informal settlements, there are no administrative lists that could be used to develop the sample frame. Instead, there was extensive engagement with local stakeholders to ensure that we found all providers operating in the area. In instances where there were additional daycare providers found through the workshops (see "Workshop" section above), we reconciled data to ensure that new providers were in fact within the community boundary and that they did meet the study criteria.

[8] Four communities did not have a sufficient number of eligible daycare providers but were otherwise suitable. We randomized these 4 communities to treatment or control groups within their own strata, and will follow up with them if our partner decides to revise downward the minimum number of identified providers.

[9] In collaboration with our partner, we also then re-visited a subset of communities to re-draw the polygon map boundary, and conducted a second round of boundary mapping and listing to ensure clear coordination and that each community fit the study criteria.

Figure 3: Number of study communities, by county



3.4.2 Data quality checks

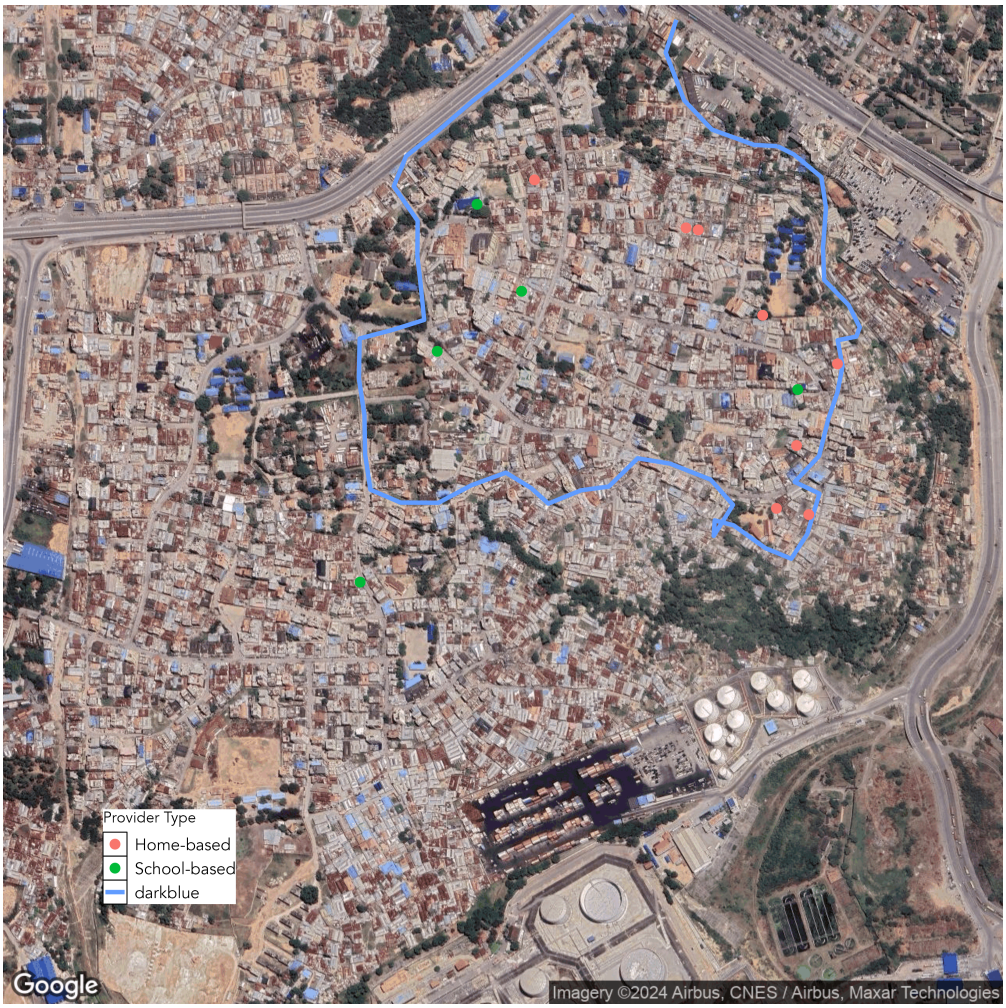
Innovations for Poverty Action-Kenya (IPA) will implement all survey rounds. During the baseline, they implemented four safeguards to ensure data quality, and they will continue these practices during subsequent rounds. First, IPA conducted real-time monitoring of data collection progress. Second, they ran high-frequency data checks to identify potential data problems, such as unexpected missing values. Third, staff conducted backchecks over the phone with a random 10% sample. And finally, staff conducted audio audits for all household phone surveys (with respondent consent) based on short (1–5 minute) recordings. The RA and field supervisors held weekly debriefings to provide feedback based on the data quality monitoring.

3.5 Baseline: Firms

Firm listing

Following the boundary mapping, we conducted a longer firm listing in February 2024. This listing also serves as the sample frame for the firm baseline. Through this activity, we captured basic characteristics of each daycare provider in the mapped areas, including business and owner name, contact information, GPS coordinates, type of provider, and number of children served by age group. Through this activity, we found 1,222 daycare providers eligible for our baseline survey.

Figure 4: Sample community boundary and firms



Note: Coordinates jittered to mask firm locations

Table 1: Firm Sample Frame and Baseline

	Total	Control	Treatment
Listed	1,222	562	660
Attempted	1,106	510	596
Completed	978	453	525
Response rate	88%	89%	88%

Note: Listing was conducted prior to randomization. The observations in the control and treatment groups for listing represent the number of firms listed in control and treatment communities, respectively.

Firm baseline survey

Our firm baseline was conducted in person, using the firm listing as a sample frame. We randomly selected a sample of firms to participate in the baseline firm survey, ordering the remainder to serve as a backup in the event of refusals or non-response and excluding those for which we did not have phone numbers.[10] To ensure a minimum number of surveys per community, we enacted the following decision rule for inclusion in the study: In communities with 20 firms or fewer, we included all firms in the baseline; in communities with more than 20 firms, firms were randomly selected, with the number proportional to the total number of firms identified in the community to arrive at our minimum sample size.

Of the 1,222 eligible firms, we ultimately attempted to reach 1,106 firms, after dropping those without contact information and dipping into the backup list as needed for different communities. We completed surveys with 978 firms, for an overall response rate of 88%. Of the firms we attempted to but could not survey, 5% could not be reached and 5.2% declined to participate.

The 978 baseline firms were visited in March–April 2024. The baseline survey interviewed the primary owner or school principal, and captured a rich set of details on costs, revenues, and firm operations. At school-based centers this individual is not typically the person who interacts daily with the children. Thus, at school-based centers, we also interviewed the teacher for the youngest grade served. Overall, our survey includes information on enrollment numbers, ages served, revenue and costs; socio-economic variables of the teacher; and the infrastructure and features of the classroom, care provided, and feeding program availability.

The final sample of firms with complete information is 978.

3.6 Baseline: Households

Listing activity for household recruitment

Recruiting a random sample of households that currently uses or may eventually rely upon paid daycare is difficult in this context. For example, urban informal settlements feature complex and varied housing structures with poor infrastructure and paths that make it difficult to navigate,

[10] Even though this survey was in person, we excluded those without phone numbers with our midline and endline in mind, as these firms and firm owners will be harder to locate a year or two later.

making traditional canvassing approaches (such as a random walk) error-prone. In our context additionally, caregivers who are home during any canvassing activity are disproportionately likely to not have children in daycare, while caretakers that rely upon paid daycare are typically not at home because they are at work. Finally, the population of urban informal settlements tends to be highly mobile. Due to these challenges, we elected to do a convenience sample of households with at least one child aged 3 or under living in the community as defined by the boundary exercise. Households were recruited within the mapped communities in early April 2024 through two methods: (1) recruitment outside daycare centers in our baseline firm survey and (2) approaching individuals who are around the community.

To recruit households who are highly likely to have children exposed to the intervention, we focused on households who were using paid daycare at baseline. Following the firm baseline survey, enumerators asked the firm survey respondent for permission to recruit parents with children age 3 and under at daycare drop-off or pick-up. If permitted, enumerators approached parents, recruiting those who have a child born in 2021—2024. Parents with a child older than 3 in the daycare and a younger, eligible child who is not enrolled are included in this recruitment activity.

Separately, households were also recruited from the community through visits to homes and markets following the boundary maps. Adults in the community were invited to participate if they have a child born in 2021–2024 living in the household, regardless of whether they use daycare services.

Among the completed survey respondents, enumerators collected their name, detailed contact information, and basic information on their use of daycare services. We listed 5,396 households—2,847 daycare users and 2,549 non-daycare users—across our 51 communities.

Household baseline survey

We randomly selected households from the household recruitment listing activity for the baseline survey, stratifying by daycare usage. We targeted 3,000 households in the baseline survey, or approximately 52 per community, providing a backup list to account for refusals or inability to reach the household by phone. Through phone calls, we interviewed the primary caregiver (self-identified) of a child under age 3 in the households in June and July 2024. The survey includes questions on the socio-economic status of the household, labor force participation, time use of the primary caregiver (for a child under age 3 in the household), and daycare use and expenses. We surveyed 2,820 households out of 3,624 attempted across the 51 study communities, with a 77.8% response rate. This response rate is balanced between treatment and control communities (see Table 2). The final sample of households for our baseline analysis is 2,820.

Table 2: Household Sample Frame and Baseline

	Total	Control	Treatment
Listed	5,396	2,659	2,737
Attempted	3,624	1,805	1,819
Completed	2,820	1,393	1,427
Response rate	77.8%	77.1%	78.4%

Note: Listing was conducted prior to randomization. The observations in the control and treatment groups for listing represent the number of households listed in control and treatment communities, respectively.

3.7 Randomization

In April 2024, after the conclusion of the firm survey we randomly assigned the study communities into treatment or control groups, stratifying at the county level.[11]

Table 3 presents the balance between these two groups at the firm level. It demonstrates that randomization created comparable groups as measured by the majority of characteristics. In both treatment and control groups, 85% of owners are female, and the mean age is 43–44 years old. There is a slight imbalance in the percent of owners that report secondary school as their highest degree ($p = 0.062$), although the share that has completed some post-secondary schooling is approximately equal at 77%.

Mean characteristics are similarly balanced at the firm level. Around 71% of all daycares are school-based. The mean daycare has 9–10 children present, and approximately 1.2 caregivers were present at the time of the survey; the number of caregivers is slightly imbalanced when including strata fixed effects. Profits and revenues are also balanced; because these measures are highly skewed, we use them and test for balance after they have been logged and winsorized at the 1st and 99th percentile. We create indices to capture various dimensions of quality: hygiene and safety; toys and manipulatives; and child experience, each normalized with respect to the control group, following Anderson (2008). All of these measures are balanced between treatment and control communities. For ease of interpretation, at endline we total all items within the index that are present, and then normalize with respect to the control group.

3.8 March phone survey

We will contact daycare providers and caregivers initially surveyed at baseline in early 2025 for a brief phone survey aimed to better understand the demand for childcare and demand for specific childcare amenities. This follow-up allows us to collect midline information on daycare usage at the household- and provider-level. Additionally, it helps maintain the relationship with both households and providers, promoting higher response rates for our later surveys.

[11] We pool two counties in close proximity, for 10 strata.

Table 3: Balance Table: Firm Characteristics

Variable	(1) Control N/Clusters	(1) Control Mean/(SE)	(2) Treatment N/Clusters	(2) Treatment Mean/(SE)	(3) Randomization Inference p-values
Panel A: Owner Characteristics					
Female	453	0.837	525	0.862	0.412
	25	(0.021)	25	(0.028)	
Respondent's age	453	43.256	525	44.334	0.221
	25	(1.340)	25	(0.777)	
Highest Degree: Secondary	453	0.158	525	0.109	0.062
	25	(0.031)	25	(0.023)	
Highest Degree: Post-secondary	453	0.771	525	0.773	0.947
	25	(0.049)	25	(0.047)	
Any childcare-specific training	453	0.642	525	0.694	0.269
	25	(0.067)	25	(0.043)	
Panel B: Firm Characteristics					
School-based Daycare	453	0.720	525	0.708	0.852
	25	(0.066)	25	(0.058)	
Number of children present	453	9.082	525	10.300	0.181
	25	(1.217)	25	(0.817)	
Number of caregivers present	453	1.189	525	1.275	0.054
	25	(0.063)	25	(0.036)	
PPI Score	453	66.862	525	65.878	0.530
	25	(1.738)	25	(1.433)	
Monthly profits (USD)	453	803.172	525	831.587	0.810
	25	(164.817)	25	(95.839)	
Monthly revenues (USD)	453	991.131	525	1012.252	0.880
	25	(166.049)	25	(99.666)	
Childcare Experience Index	453	0.340	524	0.254	0.536
	25	(0.160)	25	(0.124)	
Toys and Manipulatives Index	453	0.128	524	0.210	0.343
	25	(0.091)	25	(0.077)	
Hygiene and Safety Index	453	0.074	524	0.092	0.890
	25	(0.094)	25	(0.107)	
F-test of joint significance (F-stat)					0.887
F-test, number of observations					817
F-test, number of clusters					51

Notes: The observations are based on completed surveys.

3.9 Midline: Firms

We will conduct a midline, post-intervention survey of all firms tentatively scheduled for June 2025. In this data collection activity we will a) re-map the study area and determine the number of firms in each community; and b) re-visit all firms that were part of the baseline and ask them largely the same set of questions.

3.10 Endline: Firms and households

We will conduct a second, post-intervention survey of all firms tentatively scheduled for June 2026. In this data collection activity we will a) re-map the study area and determine the number of firms

Table 4: Balance Table: HH Background Characteristics

Variable	(1) Control N/Clusters	Mean/(SE)	(2) Treatment N/Clusters	Mean/(SE)	(3) Randomization Inference p-values
Panel A: Household Characteristics					
Household Head	1393	0.363	1427	0.363	0.998
	25	(0.023)	25	(0.021)	
Married	1314	0.718	1352	0.724	0.824
	25	(0.030)	25	(0.023)	
No education	1393	0.005	1427	0.003	0.701
	25	(0.005)	25	(0.005)	
Completed pre-primary	1393	0.002	1427	-0.001	0.303
	25	(0.002)	25	(0.003)	
Completed primary education	1393	0.038	1427	0.032	0.644
	25	(0.017)	25	(0.011)	
Completed junior secondary	1393	0.227	1427	0.218	0.688
	25	(0.030)	25	(0.019)	
Completed senior secondary	1393	0.415	1427	0.461	0.054
	25	(0.022)	25	(0.021)	
Completed tertiary	1393	0.314	1427	0.287	0.226
	25	(0.043)	25	(0.020)	
Child under 5 y/o	1393	1.125	1427	1.126	0.962
	25	(0.020)	25	(0.017)	
Mother is the primary caregiver	1391	0.880	1425	0.870	0.531
	25	(0.029)	25	(0.015)	
Pay for care during March 2024	1391	0.582	1425	0.564	0.307
	25	(0.023)	25	(0.016)	
Monthly childcare costs (KES)	1393	1421.843	1427	1380.521	0.802
	25	(118.121)	25	(142.694)	
Panel B: Labor Force					
Paid work in the past 7 days	1393	0.394	1427	0.380	0.547
	25	(0.030)	25	(0.019)	
Working mother	1185	0.558	1204	0.542	0.443
	25	(0.037)	25	(0.018)	
Total hours worked while providing childcare	1393	2.565	1427	2.476	0.768
	25	(0.316)	25	(0.269)	
F-test of joint significance (F-stat)					0.978
F-test, number of observations					2245
F-test, number of clusters					50

Notes: The observations are based on completed surveys. Monthly childcare costs include zeros.

in each community; and b) re-visit all firms that were part of the baseline and ask them largely the same set of questions.

During this activity we will also attempt to re-contact households that were part of the baseline survey. This activity will be done in person.

3.11 Power analysis

Using data-informed parameters, we conclude that our study is well-powered for our primary outcomes of interest. We use a significance level of 5% and estimate minimum detectable effect (MDE) sizes with 80% power. We report MDEs based on our intention-to-treat effects for our entire sample

of firms and households and for two subsamples: likely-takers and our spillover sample.[12]

Our likely-takers are the set of firms (60%) that, in both treatment and control communities, participated in the pre-intervention workshop. In the treatment areas, 82.5% of this group ultimately enrolled in the program, receiving the mentorship, grant, and continuous support. In the control group, workshop attendees are a comparison group of firms that would also be highly likely to take up the program if offered.

The remainder form the spillover sample, the set of 395 firms (39%) that were invited to but did not participate in a pre-intervention workshop. Note that the enrollment rate for this group is close to zero (0.3%)—only one firm that did not join the workshop ultimately enrolled in the program. This spillover sample is of particular interest because the direction of impacts are unknown: improved market quality through the treatment may pressure other firms in the same communities to increase their quality, or it may lead them to differentiate and offer a lower-cost, lower-quality service to households. We assume a follow-up response rate of 90%.

We expect the largest and most direct impacts of firm entry to be on daycare quality (HA1), using a quality index normalized to the control group. At baseline, the ICC of our childcare quality index is 0.096, controlling for stratification-cell fixed effects.

With 453 control firms in 26 communities and 525 treatment firms in 25 communities at baseline, we are therefore powered to detect a 0.29 s.d. change in childcare quality across all firms in treatment communities. Reflecting ITT effects, we are powered to detect a 0.36 s.d. change in childcare quality in our spillover sample and a 0.32 standard deviation change in our likely-taker sample.[13]

We believe these effect sizes are reasonable given the substantial direct quality improvements provided by our partner. We contextualize these magnitudes in Appendix Tables A6 and A7. First, in Appendix Table A6 we perform a "what-if" exercise for whether firms in treatment communities who attended the workshop (i.e., compliers) shifted into offering an amenity they did not have at baseline for each of the quality indices. All non-complier firms in treatment communities and firms in control communities are assumed to be the same. In Row 4, we analyze if all items in the index increase (i.e., they "max-out" all items). As expected, these magnitudes are substantial, as measures of childcare quality are typically low in this setting (see, for example, Appendix Table A5). Rows 5–7 show more modest changes still lead to a well-powered study in both the aggregate sample and subindices. Row 5 changes one amenity in each index; row 6 changes a different amenity; row 7 changes both amenities.[14] As long as 1–2 items per index increase on average for treatment compliers, our study has a high level of statistical power. We believe this is likely as the intervention

[12]We could think about power for compliers by doubling our MDEs (reflecting take-up of 0.5), but we note this would be biased due to the likelihood of spillovers. For the sub-sample of likely-takers, we would similarly scale by 1.2 (reflecting that 82.5% of workshop attendees enrolled.)

[13]While the ITT MDE is larger for the likely taker sample relative to the full sample, reflecting the smaller sample size, we note that power is greater because the take-up rate among this subsample is much higher, at 82.5%

[14]Row 5 changes a jug in the safety index, stacking cups in the toys index, the provider keeps a schedule in the experience index; row 6 changes a clean rug in the safety index; a doll in the toys index; whether the provider reads to children in the experience index. These were chosen based upon the structure of the intervention and what was likely to change in our opinion.

directly leads to an increase in the presence of multiple types of toys and provides multiple hygiene and safety improvements.

Second, for quantity, we rely on child-days of attendance, accounting for the substantial variation in number of days children attend and noting that childcare use might change on the intensive and extensive margin. On average, firms provide 65.3 child-days per week of care, with a standard deviation of 55.3 child-days. The likely-taker sample has a similar mean and standard deviation. Attendance is highly correlated with daycare type, as school-based providers serve far more children than home- and center-based providers. Indicators for daycare type and strata predict 14% of the variance in this outcome variable. The residualized ICC in our baseline data is 0.042. This yields a MDE of 0.22 s.d., or 12.2 child-days for the full sample, 0.27 s.d (14.8 child days) for the likely-taker sample, and 0.31 s.d (17.1 child days) for the spillover sample. In the full sample of firms, we have adequate statistical power with an increase of at least 8 child-days among treatment compliers, and excellent power with 10 child-days per week among treatment compliers. Among the likely-taker sample we have adequate statistical power so long as slightly more than 1 child-day increases among all treamtent compliers. We believe that this is reasonable, as parents may feel more comfortable leaving their children at daycare firms with quality improvements. We note that these changes could be on either the intensive or extensive margin; any extensive margin changes would likely translate into multiple child-days per firm. Finally, in our interviews with providers and discussions with our partners, providers are unlikely to be capacity-constrained, and we will measure this quantitatively during our March phone survey.

Because revenue is noisy, we measure provider revenue through three follow-up surveys (March phone follow-up, midline, and endline) and calculate power by simulation. We assume our specification will be ANCOVA, an approach that will increase statistical power (McKenzie, 2012). We assume revenue is autocorrelated with $\rho = 0.35$ between periods, in line with the average autocorrelation of small enterprise profits (McKenzie, 2012). For revenue, the baseline mean is \$1,036 with a standard deviation of \$1,306, and ICC of 0.04. Estimating power based on pooling our two rounds of follow-up data, assuming 10% pre-round attrition, we are powered for a minimum detectable effect size, based on ITT estimates, of \$350 across all firms, or 34%. This figure is smaller than three of the firm training papers cited in the McKenzie (2021) literature review with positive effects. We also note that our sample size, with 525 firms in treatment, is on par or larger than most of the microenterprise studies included in the meta-analysis.

Within the spillover sample (those who did not attend the workshop), we are powered to detect a change of \$335 (34% relative to the baseline average revenue of \$977 for this group).[15] Within the likely-taker sample, we are powered to detect an increase of \$475 (44%) relative to a baseline mean of \$1,077 for this group. This MDE is again smaller than the effects identified in three studies looking at training that find positive effects (McKenzie, 2021), and the training and intervention in the current study are arguably much more intense than those included in the meta-analysis.

[15]The MDE in dollars decreases relative to the full sample because standard deviation and ICC are smaller for this group.

The current intervention involves months of continuous training and benefits – during the 3-month training but also once firms become franchisees, – a large quality improvement grant of USD 200, and regular porridge deliveries (facilitating improved quality for daycares not offering food at baseline). Moreover, these quality improvements are directly observable to potential clients, so may have a more direct effect on revenue, beyond a provider's increased ability to run a business.

4 Empirical analysis

4.1 Econometric specification and outcomes

In this section, we present our econometric specification and discuss our primary and secondary outcomes of interest.

We measure the intention-to-treat effect of community-level assignment on our outcomes of interest using the following specification:

$$\mathrm{Y}_{ic} = \alpha + \beta_1 \mathrm{Treatment}_c + Y_{ic}^{BL} + \phi_c + \epsilon_{ic} \tag{1}$$

where Y_{ic} is the outcome of interest for respondent i in community c. $Treatment$ is an indicator variable indicating the randomly assigned treatment status of the community. We include, when available, the baseline measure of the outcome variable, Y_{ic}^{BL} , and a set of strata (county) fixed effects, ϕ_c. Although we present for brevity the estimating equation for a given follow-up wave, we will also examine both follow-up rounds combined together in the panel dataset for all variables of interest. All standard errors will be clustered at the community level, our unit of randomization.

To improve the precision of our estimates, we will also estimate a version that includes other baseline covariates chosen following a double-selection lasso procedure, following Urminsky et al. (2016). If this specification meaningfully improves our precision, we will use these as our main results and present the version without additional covariates in the appendix.

Additionally, we will also estimate an alternative specification in which we pool observations over multiple follow-up rounds, as described in our discussion of statistical power.

4.1.1 Likely taker and spillover analysis

Participation in the workshop is a very strong predictor of enrollment in the program: 82.5% of attendees in treatment communities enrolled in the program, while only one non-attendee (0.3%) similarly enrolled. As noted above, we find that participation rates and baseline characteristics are balanced between treatment and control communities (see Appendix Tables A1 and A2). We therefore use workshop participation to divide our sample into the set of "likely takers" who attended the workshop and "spillover firms" who did not attend the workshop.

We will also estimate equation (1) for each sub-sample. As explained in more detail in Section 3.11, the higher take-up rate among likely takers increases our statistical power, while the spillover sample will allow us to directly test for within-market spillovers.

4.1.2 Multiple hypothesis testing

To avoid over-rejecting our null hypotheses as a result of multiple hypothesis testing, we will report MHT-adjusted p-values for our secondary outcomes using the Westfall-Young free step-down resampling method (Westfall and Young, 1993) alongside unadjusted standard errors. We will adjust within families of outcomes for these secondary hypotheses (where families are described above).

4.1.3 Missing baseline variables

For any baseline variables that are missing due to refusal, non-response, or respondent indicating "don't know," we will recode missing values as zero and include missing value flags in our analysis.

4.2 Additional outcomes

We also plan to include a set of other outcomes that are relevant for families and/or policymakers looking to improve the quality of daycare in the private sector.

- **Labor supply of spouse/ partner** of primary caregiver
- **Wages of the primary caregiver**, calculated as the earnings over the past 30 days divided by total hours worked during the same period (Questions F3, F4, F7, and F8).
- **Household income** over the past 30 days (Question C05).
- **Community-level daycare closures**. We will measure whether the presence of the partner organization led to entry or exit of daycare firms in the community.
- **Community-level non-franchise enrollment**. We will measure whether the presence of the partner organization affected enrollment in non-franchise centers in the community.
- **Other Measures of Quality**. We will also explore other measures of childcare quality that are used in developed-country contexts and that are important from a policy perspective. However, we view these as exploratory as the intervention may not change them and/or our study design is not set up to measure them well. These measures include: the caregiver-to-child ratio; an age-adjusted caregiver-to-child ratio; enumerator observations regarding whether or not the provider was positively engaging with children (i.e., playing or holding/comforting) during the time of the visit, as well as use of handwashing facilities and diapering; and whether providers give medicine to children without parental consent.

4.3 Heterogeneous effects

4.3.1 Firm-level analysis

At the firm level, we anticipate the effectiveness of the intervention may depend on the initial quality of the enrolling firm and the number of children served. School-based centers tend to be

larger and of higher quality, but they also differ in their management structure and characteristics of providers. We plan to test for treatment heterogeneity based on the following three dimensions:

1. (1) school-based or non-school-based (i.e., home or center-based) daycare (Question A16);
2. (2) the number of children served, split at the overall median (Questions C02 and Q1); and
3. (3) the overall firm quality index, split at the baseline community-level median.

4.3.2 Household-level analysis

Among households, we anticipate that those with any children in daycare may be most likely to be directly affected by improved daycare quality; they would receive the largest "dosage" if their daycare becomes a franchisee. Alternatively, the opposite dynamic could happen: if improvements in quality lead non-using households to be more likely to use the new high-quality care, we may see the largest impacts on mothers' labor supply and earnings among this group.

Similarly, child age at baseline will affect their responsiveness to improved quality, with younger children potentially benefiting most, but it will also affect the period of exposure before they proceed to primary school. For this reason, we plan to test for household-level treatment heterogeneity based on (a) whether their child born in 2021–2024 was attending daycare at baseline (Questions D04 and D05); (b) whether the mother was working for pay at baseline (Questions F1 and G2); and (c) child age (Question C04). We also plan to test for heterogeneity based on (d) baseline Household PPI score, above/below the community-level median.

5 Limitations and challenges

There are several limitations and challenges to the completion of this study. First, residents of informal settlements may be reluctant to trust outsiders due to sporadic, unplanned raids and demolitions. Our risk mitigation strategy was to extensively liaise and engage with local stakeholders and policymakers to maintain trust and buy-in. Relatedly, we also acknowledge risks from unplanned closures of daycares from specific communities due to sporadic crack-downs. As in these events, daycares typically reopen after some time, we plan to adjust field operations to minimize study attrition.

Secondly, residents living in urban informal settlements are disproportionately likely to be in poverty and may be unable to engage fully in the intervention. Thus, there may be substantial opportunity costs of time and money leading to insufficient take-up and completion of the intervention. We have attempted to minimize this risk by completing a workshop in all communities (treatment and control) to identify those providers with a high propensity to participate in the intervention.

Finally, due to the ever-changing political situation in Kenya, there may be protests that may delay field activities, or require adaptation from the plan. Protests may unexpectedly turn violent and thus data collection needs to be carefully executed to ensure enumerator safety. Now that the

baseline survey frame is completed, the risk mitigation strategy will be to shorten the midline and endline surveys and conduct surveys by phone.

6 Administrative information

6.1 Funding

This work is funded by the UK Foreign, Commonwealth & Development Office (FCDO) and the Institute of Labor Economics (IZA) within the G²LM||LIC Programme. We are grateful for the additional financial support provided by the Bill and Melinda Gates Foundation and the Livelihood Impact Fund (LIF). Dr. Fitzpatrick gratefully acknowledges support from J-PAL Gender and Economic Agency Initiative that funded large-scale pilot work related to this project.

6.2 Ethical review

The Strathmore University Institutional Scientific and Ethical Review Committee (SU-ISERC) has provided ethical review for this study, initial approval SU-ISERC1602/23. This study has also received IRB approval from The Ohio State University (submission #2023B0300).

6.3 Declarations of interest

Declaration of interest: None.

6.4 Acknowledgements

This study would not have been possible without excellent research assistance from Miryana Vinka Dayanti. We thank Eric Ochieng, Eunice Kioko and Dorine Mudaki, and the entire IPA-Kenya team for their tireless efforts. Finally, we thank all of our respondents for their time and trust, as well as our partner organization for their ongoing support and cooperation.

References

Ajayi, K. F., A. Dao, and E. Koussoubé (2022). *The Effects of Childcare on Women and Children*. Washington, DC: World Bank. Working Paper.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association 103*(484), 1481–1495.

Anderson, S. J. and D. McKenzie (2020). Improving business practices and the boundary of the entrepreneur: A randomized experiment comparing training, consulting, insourcing, and outsourcing. *World Bank Policy Research Working Paper Series* (9502).

Andrabi, T., J. Das, and A. I. Khwaja (2017). Report cards: The impact of providing school and child test scores on educational markets. *American Economic Review 107*(6), 1535–1563.

Andrew, A., O. P. Attanasio, R. Bernal, L. C. Sosa, S. Krutikova, and M. Rubio-Codina (2024). Preschool quality and child development. *Journal of Political Economy 132*(7), 000–000.

Baker, M., J. Gruber, and K. Milligan (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy 116*(4), 709–745.

Bakhtiar, M. M., G. Bastian, and M. Goldstein (2022). Business training and mentoring: Experimental evidence from women-owned microenterprises in ethiopia. *Economic Development and Cultural Change 71*(1), 151–183.

Bartik, T. J. (2022). *The long-run effects of high-quality pre-K: What does the research show?* Upjohn Institute Working Paper.

Bjorvatn, K., D. Ferris, S. Gulesci, A. Nasgowitz, V. Somville, and L. Vandewalle (2022, April). Childcare, Labor Supply, and Business Development: Experimental Evidence from Uganda.

Blau, D. M. (2001). *Child care problem: An economic analysis*. Russell Sage Foundation.

Brooks, W., K. Donovan, and T. R. Johnson (2018). Mentors or teachers? microenterprise training in kenya. *American Economic Journal: Applied Economics 10*(4), 196–221.

Bruhn, M., D. Karlan, and A. Schoar (2018). The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in mexico. *Journal of Political Economy 126*(2), 635–687.

Busso, M. and S. Galiani (2019). The causal effect of competition on prices and quality: Evidence from a field experiment. *American Economic Journal: Applied Economics 11*(1), 33–56.

Carneiro, P., J. Das, and H. Reis (2024). The value of private schools: Evidence from pakistan. *Review of Economics and Statistics 106*(5), 1301–1318.

Cascio, E. U. (2009). Maternal labor supply and the introduction of kindergartens into american public schools. *Journal of Human resources 44*(1), 140–170.

Clark, S., C. W. Kabiru, S. Laszlo, and S. Muthuri (2019). The impact of childcare on poor urban women's economic empowerment in africa. *Demography 56*(4), 1247–1272.

Cohodes, S. R. and J. J. Feigenbaum (2021). Why does education increase voting? evidence from boston's charter schools. *National Bureau of Economic Research Working Paper* (29308).

De Mel, S., D. McKenzie, and C. Woodruff (2013). What generates growth in microenterprises? experimental evidence on capital, labor and training. Technical report, Mimeo, World Bank.

Delecourt, S. and A. Fitzpatrick (2021). Childcare matters: Female business owners and the baby-profit gap. *Management Science 67*(7), 4455–4474.

Delecourt, S., A. Marchenko, A. Fitzpatrick, and L. Lowe (2022). Too close to home: Women's business location and the gender profit gap.

Devercelli, A. E. and F. Beaton-Day (2020). Better jobs and brighter futures. Technical report, World Bank, Washington, DC.

Donald, A., S. Lowes, and J. Vaillant (2023). Experimental evidence on rural childcare provision.

Fiala, N. (2013). Stimulating microenterprise growth: Results from a loans, grants and training experiment in uganda. Technical report, Working Paper.

Gaynor, M. and R. J. Town (2011). Competition in health care markets. *Handbook of health economics 2*, 499–637.

Gray-Lobe, G., P. A. Pathak, and C. R. Walters (2021). *The long-term effects of universal preschool in Boston*. Number w28756. National Bureau of Economic Research.

Heckman, J. J. and G. Karapakula (2019). *Intergenerational and intragenerational externalities of the Perry Preschool Program*. Number 25889. NBER Working Paper.

Hotz, V. J. and M. Xiao (2011). The impact of regulations on the supply and quality of care in child care markets. *American Economic Review 101*(5), 1775–1805.

Hughes, R. C., R. Muendo, S. S. Bhopal, S. Onyango, E. W. Kimani-Murage, B. R. Kirkwood, Z. Hill, and P. Kitsao-Wekulo (2023). Parental experiences of childcare in an informal urban settlement: qualitative interview findings from the nairobi early childcare in slums (necs) project. *BMJ open 13*(4), e071627.

Jakiela, P., O. Ozier, L. C. Fernald, and H. A. Knauer (2024). Preprimary education and early childhood development: Evidence from government schools in rural kenya. *Journal of Development Economics*, 103337.

Jayachandran, S. (2021). Microentrepreneurship in developing countries. *Handbook of Labor, Human Resources and Population Economics*, 1–31.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics 99*(2), 210–221.

McKenzie, D. (2021). Small business training to improve management practices in developing countries: re-assessing the evidence for 'training doesn't work'. *Oxford Review of Economic Policy 37*(2), 276–301.

McKenzie, D. and S. Puerto (2021, April). Growing Markets through Business Training for Female Entrepreneurs: A Market-Level Randomized Experiment in Kenya. *American Economic Journal: Applied Economics 13*(2), 297–332.

McKenzie, D., C. Woodruff, K. Bjorvatn, M. Bruhn, J. Cai, J. Gonzalez-Uribe, S. Quinn, T. Sonobe, and M. Valdivia (2021). Training entrepreneurs. *VoxDevLit 1*(2), 3.

Neuman, M. J. and S. Powers (2021). Political prioritization of early childhood education in low-and middle-income countries. *International Journal of Educational Development 86*, 102458.

OANDA (2024). Average exchange rates (march-may 2024): Usd to kes. Accessed: 2024-07-09.

Quinn, S. and C. Woodruff (2019). Experiments and entrepreneurship in developing countries. *Annual Review of Economics 11*(1), 225–248.

Ranganathan, A. and D. Pedulla (2018). *Organizations as network equalizers? Employer-provided childcare and the labor supply of working mothers.* Working paper.

Urminsky, O., C. Hansen, and V. Chernozhukov (2016). Using double-lasso regression for principled variable selection. *Available at SSRN 2733374*.

Westfall, P. H. and S. S. Young (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*, Volume 279. John Wiley & Sons.

Appendix

Table A1: Characteristics of those who did not attend workshop, treatment and control communities

	(1) Control		(2) Treatment		(3) Randomization Inference
Variable	N/Clusters	Mean/(SE)	N/Clusters	Mean/(SE)	p-values
Panel A: Owner Characteristics					
Female	189	0.816	206	0.859	0.222
	25	(0.041)	25	(0.034)	
Respondent's age	189	43.645	206	44.478	0.512
	25	(2.151)	25	(1.138)	
Highest Degree: Secondary	189	0.197	206	0.166	0.435
	25	(0.050)	25	(0.036)	
Highest Degree: Post-secondary	189	0.710	206	0.691	0.802
	25	(0.069)	25	(0.065)	
Any childcare-specific training	189	0.563	206	0.625	0.285
	25	(0.069)	25	(0.049)	
Panel B: Firm Characteristics					
School-based Daycare	189	0.638	206	0.598	0.682
	25	(0.090)	25	(0.080)	
Number of children present	189	9.590	206	9.542	0.974
	25	(1.734)	25	(1.144)	
Number of caregivers present	189	1.117	206	1.253	0.002
	25	(0.040)	25	(0.034)	
PPI Score	189	66.637	206	65.716	0.656
	25	(1.875)	25	(1.907)	
Monthly profits (USD)	189	980.846	206	925.841	0.690
	25	(222.158)	25	(103.391)	
Monthly revenues (USD)	189	1101.901	206	1050.315	0.730
	25	(240.403)	25	(113.319)	
Childcare Experience Index	189	-0.027	205	0.037	0.662
	25	(0.207)	25	(0.141)	
Toys and Manipulatives Index	189	-0.007	205	0.138	0.215
	25	(0.132)	25	(0.107)	
Hygiene and Safety Index	189	0.089	205	0.365	0.068
	25	(0.125)	25	(0.135)	
F-test of joint significance (F-stat)					0.967
F-test, number of observations					313
F-test, number of clusters					50

Note: Observations limited to firms that did not attend a half-day workshop held by our partner organization. Based on completed baseline surveys.

Table A2: Characteristics of those who did attend workshop, treatment and control communities

Variable	(1) Control N/Clusters	(1) Control Mean/(SE)	(2) Treatment N/Clusters	(2) Treatment Mean/(SE)	(3) Randomization Inference p-values
Panel A: Owner Characteristics					
Female	264	0.850	319	0.859	0.843
	25	(0.031)	25	(0.043)	
Respondent's age	264	43.041	319	44.206	0.261
	25	(0.933)	25	(0.933)	
Highest Degree: Secondary	264	0.119	319	0.059	0.050
	25	(0.031)	25	(0.030)	
Highest Degree: Post-secondary	264	0.832	319	0.842	0.876
	25	(0.031)	25	(0.053)	
Any childcare-specific training	264	0.706	319	0.751	0.434
	25	(0.080)	25	(0.052)	
Panel B: Firm Characteristics					
School-based Daycare	264	0.800	319	0.800	0.996
	25	(0.047)	25	(0.068)	
Number of children present	264	9.091	319	11.047	0.079
	25	(1.392)	25	(0.930)	
Number of caregivers present	264	1.238	319	1.283	0.412
	25	(0.091)	25	(0.047)	
PPI Score	264	67.104	319	65.922	0.462
	25	(2.080)	25	(1.470)	
Monthly profits (USD)	264	712.848	319	727.074	0.927
	25	(161.945)	25	(162.856)	
Monthly revenues (USD)	264	950.243	319	956.958	0.968
	25	(193.359)	25	(166.462)	
Childcare Experience Index	264	0.589	319	0.430	0.351
	25	(0.102)	25	(0.154)	
Toys and Manipulatives Index	264	0.225	319	0.257	0.731
	25	(0.087)	25	(0.091)	
Hygiene and Safety Index	264	-0.011	319	-0.177	0.197
	25	(0.111)	25	(0.119)	
F-test of joint significance (F-stat)					0.910
F-test, number of observations					504
F-test, number of clusters					50

Note: Observations limited to firms that attended a half-day workshop held by our partner organization. Based on completed baseline surveys.

Table A3: Data Appendix: Primary Hypotheses

Primary Hypothesis	**Data Source**
HA1. The social franchising model will improve daycare quality	• **Hygiene and Safety Index:** Which of the following are present in the facility? (check all that apply, ask if not obvious): basin/sink for handwashing, jug/bottle for handwashing, potties for toilet training, changing table/ diapering area, mattress or mats for napping, first-aid-kid or box; Business appeared clean (strongly disagree, disagree, agree, strongly agree; agree or strongly agree=1); Business appeared safe (strongly disagree, disagree, agree, strongly agree; agree or strongly agree=1); clean rug was on the ground • **Toys and Manipulatives Index:** (If toys checked) What kind of toys do you have for the children?: Cups for playing/stacking, blocks for building/stacking, vehicles/rolling items, dolls, rattles/shakers, musical instruments, costumes/dress up, art supplies (paint, crayons, etc), string/laces for stringing beads. • **Child Experience Index:** Do the children typically follow a daily schedule?: yes/no; Think about the yesterday or the last day you provided care. Did you sing songs with the children?: yes/no; Think about the yesterday or the last day you provided care. Did you read to the children?; (Do not ask) Is the schedule/time table of the day posted on the wall?; Which of the following are present in the facility? (Check all that apply. Ask if not obvious): decorations on walls; Is there an outside area that children can use for playing?
HA2. The social franchising model will increase provider revenue	• Number of children • Days per week attended • Daily rate

Continued on next page

Primary Hypothesis	**Data Source**
HA3. The social franchising model will increase child-days served	• Number of children • Days per week attended

Table A4: Data Appendix: Secondary Hypotheses

Secondary Hypothesis	**Data Source**
HB1. The social franchising model will have spillover effects onto competing providers	Total number of children that have attended at least 1 day in the past 7 days, Number of children by age group • What is the total number of children that have attended at least 1 day in the past 7 days? (Include both daycare and part-time/ after school attendees.) • How many children in each age group attended at least 1 day in the past 7 days? (Tick all that apply): $<$ 6 months, 6 months – 1 year, 1 – 2 years, 2 years, 3 years, 4 years, 5 years, 6 years and older
HB2. The social franchising model will increase provider profits	Not explicitly listed; use revenue (HA3) and costs

Continued on next page

Secondary Hypothesis	**Data Source**
HB3. Social franchising model will increase provider agency and pride in profession	Agency and pride in profession, self-efficacy, and social standing • What do you think is a more respected profession in your community: styling hair or running a daycare? (D03) • What do you think is a more respected profession in your community: selling chapati or [answer to D03]? • I feel like I can solve most problems that I face in my day-to-day life: Strongly agree, agree, neutral, disagree, or strongly disagree • Other people give me the respect that I deserve: Strongly agree, agree, neutral, disagree, or strongly disagree • Now, assume that the top rung of the ladder (10) represents the best-off members of your community and that the lowest rung (1) are the poorest individuals. Where would you place your household on the ladder in terms of economic status?: between 1-10 or don't know
HB4. The social franchising model will increase household usage of childcare	Household's daycare use at the extensive and extensive margin Household's daycare use, extensive margin: • During [previous month], did you or anyone in your household ever pay anyone money to take care of [CHILD NAME]? • Which sources of paid care did you use to care for [CHILD NAME] during [previous month]? Household daycare use, intensive margin: • About how many days per week did [CHILD NAME] spend at [daycare/school] during a typical week in [previous month]?

Continued on next page

Secondary Hypothesis	**Data Source**
HB5. The social franchising model will increase maternal employment and household earnings	Primary caregivers' labor supply extensive and intensive margin • Primary caregivers' labor supply, extensive margin: In the past 7 days did you work for at least one hour as an employee for wage, salary, commission, including doing paid domestic work or farm work not on your own farm?; In the past 7 days did you work for at least one hour on a business or income-generating activity that you own or operate? Do not include own farm labor. (Can clarify: For example, as a trader, shopkeeper, barber, dressmaker, carpenter, taxi driver, car washer, etc.?); In the past 7 days did you help (for at least one hour) in any non-farm business enterprise belonging to or run by someone else in your household? • Primary caregivers' labor supply, intensive margin: How many total hours did you work in the last 7 days as an employee?, How many overall hours did you work in the last 7 days at your own business?, How many overall hours did you work on this enterprise or enterprises?
HB6. The social franchising model will increase select measures of child development	Child development outcomes based on caregiver reported-responses. We will file an amendment to the pre-analysis plan once we have piloted this assessment instrument, which will be employed at the endline, in June 2026.
HB7. The social franchising model will improve child nutritional intake through increased daycare feeding	• What meals, if any, do you provide to any children at this daycare? • How often in the past month... a child arrives without food and is hungry during the day.

Table A5: Select Quality Information

Variable	(1) Control, Do not Enter N/Clusters	Mean/(SE)	(2) Treatment, Kidogo Enters N/Clusters	Mean/(SE)	(1)-(2) Pairwise t-test N/Clusters	Mean difference
Hygiene and Safety						
Basin/sink for handwashing	453	0.764	524	0.779	977	-0.015
	25	(0.031)	26	(0.039)	51	
Jug/bottle for handwashing	453	0.481	524	0.487	977	-0.005
	25	(0.042)	26	(0.053)	51	
Potties for toilet training	453	0.408	524	0.485	977	-0.076
	25	(0.042)	26	(0.044)	51	
Toys and Manipulatives						
Any toys	453	0.552	524	0.557	977	-0.005
	25	(0.032)	26	(0.037)	51	
Any books	453	0.629	524	0.626	977	0.003
	25	(0.040)	26	(0.046)	51	
Cups for playing/stacking	453	0.099	525	0.101	978	-0.002
	25	(0.015)	26	(0.016)	51	
Blocks for building/stacking	453	0.174	525	0.168	978	0.007
	25	(0.021)	26	(0.021)	51	
Vehicles/rolling items	453	0.318	525	0.322	978	-0.004
	25	(0.035)	26	(0.030)	51	
Art supplies (paint, crayons, etc)	453	0.230	525	0.238	978	-0.009
	25	(0.034)	26	(0.036)	51	
Child Experience						
Follow daily schedule	453	0.724	524	0.763	977	-0.039
	25	(0.037)	26	(0.040)	51	
Sing songs previous day you provided care	453	0.932	524	0.924	977	0.008
	25	(0.014)	26	(0.019)	51	
Read previous day you provided care	453	0.711	524	0.716	977	-0.005
	25	(0.037)	26	(0.051)	51	
Decorations of walls	453	0.618	524	0.611	977	0.007
	25	(0.043)	26	(0.051)	51	
Outside area used for playing	382	0.963	478	0.939	860	0.024
	25	(0.012)	26	(0.016)	51	
(HB7) Child go hungry at least once a week	453	0.375	525	0.438	978	-0.063
	25	(0.036)	26	(0.026)	51	

Note: This table presents a selection of baseline indicators of quality.

Table A6: Translation of the daycare quality MDE into intervention changes

	Safety Index (10 items)	**Toys Index (10 items)**	**Child Experience Index (6 items)**	**Overall Index (26 items)**
Panel A: Baseline Statistics and Power	(1)	(2)	(3)	(4)
1. Baseline Mean (Count of total items)	3.95	1.23	2.94	8.12
2. Baseline Std Dev (Count of total items)	2.83	1.99	2.29	6.07
3. Current MDE in ITT (SD)	0.29	0.29	0.29	0.29
Panel B: What-if Simulations				
4. Index Treatment Mean (Increase All Amenities by 1 for T compliers, in SD)	1.303	2.081	0.580	1.888
5. Index Treatment Mean (Increase 1 Amenity by 1 for T compliers, in SD)	0.175	0.246	0.088	0.241
6. Index Treatment Mean (Increase Another Amenity by 1 for T-compliers, in SD)	0.219	0.183	0.101	0.232
7. Index Treatment Mean (Increase Amenities by 2 for T-compliers, in SD)	0.375	0.406	0.172	0.446

Notes: Panel A summarizes data from the baseline in rows 1 and 2; row 3 is based upon the power calculations as described in the text. Panel B does simulations and reports how that change would affect the corresponding standardized difference between treatment and control communities, where the control mean is 0 with standard deviation 1. Row 4 adds all amenities within a given index to compliers in the treatment group; Row 5 adds 1 specific amenity within a given index to compliers in the treatment group; Row 6 adds a different amenity within a given index to compliers in the treatment group; Row 7 is both amenities within a given index for compliers in the treatment group.

Table A7: Translation of the daycare quantiity MDE into intervention changes

Panel A: Baseline Statistics and Power	**Full Sample** (1)	**Likely-Taker Sample** (2)
1. Baseline Mean (Number of child-days per week)	65.275	68.422
2. Baseline Std Dev (Number of child days per week)	55.340	56.197
3. Current MDE in ITT (SD)	0.221	0.267
Panel B: What-if Simulations		
4. Index Treatment Mean (Increase by 1 child-days for T compliers, in SD)	0.141	0.264
5. Index Treatment Mean (Increase by 4 child-days for T compliers, in SD)	0.175	0.320
6. Index Treatment Mean (Increase by 6 child-days for T compliers, in SD)	0.198	0.358
7. Index Treatment Mean (Increase by 8 child-days for T compliers, in SD)	0.221	0.396
8. Index Treatment Mean (Increase by 10 child-days for T compliers, in SD)	0.244	0.433

Notes: Panel A summarizes data from the baseline in rows 1 and 2; row 3 is based upon the power calculations as described in the text. Panel B does simulations and reports how that change would affect the corresponding standardized difference between treatment and control communities, where the control mean is 0 with standard deviation 1. Column 1 presents averages for the full sample, and Column 2 presents averages for the likely-taker sample. Row 4 adds 1 child-day for all compliers in the treatment group; Row 5 adds 4 child-days for all compliers in the treatment group; Row 6 adds 4 child-days for all compliers in the treatment group; Row 7 adds 6 child-days for all compliers in the treatment group. Row 8 adds 10 child-days for all compliers in the treatment group.